



CT/AU98/01023
09/581286

REC'D 12 JAN 1999

WIPO

PCT

Patent Office
Canberra

5

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES,
hereby certify that the annexed is a true copy of the Provisional specification in
connection with Application No. PP 4917 for a patent by CSL LIMITED filed on
29 July 1998.

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

WITNESS my hand this Fifth
day of January 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND
SALES





AUSTRALIA

Patents Act 1990

CSL LIMITED

PROVISIONAL SPECIFICATION

Invention Title:

P. gingivalis nucleotides and polypeptides

The invention is described in the following statement:

P. gingivalis nucleotides and polypeptides

FIELD OF THE INVENTION

5 The present invention relates to *P. gingivalis* nucleotide sequences, *P. gingivalis* polypeptides and probes for detection of *P. gingivalis*.

BACKGROUND OF THE INVENTION

10 Periodontal diseases are bacterial-associated inflammatory diseases of the supporting tissues of the teeth and range from the relatively mild form of gingivitis, the non-specific, reversible inflammation of gingival tissue to the more aggressive forms of periodontitis which are characterised by the destruction of the tooth's supporting structures. Periodontitis is associated
15 with a subgingival infection of a consortium of specific Gram-negative bacteria that leads to the destruction of the periodontium and is a major public health problem. One bacterium that has attracted considerable interest is *P. gingivalis* as the recovery of this microorganism from adult periodontitis lesions can be up to 50% of the subgingival anaerobically
20 cultivable flora, whereas *P. gingivalis* is rarely recovered, and then in low numbers, from healthy sites. A proportional increase in the level of *P. gingivalis* in subgingival plaque has been associated with an increased severity of periodontitis and eradication of the microorganism from the cultivable subgingival microbial population is accompanied by resolution of
25 the disease. The progression of periodontitis lesions in non-human primates has been demonstrated with the subgingival implantation of *P. gingivalis*. These findings in both animals and humans suggest a major role for *P. gingivalis* in the development of adult periodontitis.

P. gingivalis is a black-pigmented, anaerobic, asaccharolytic, proteolytic Gram-negative rod that obtains energy from the metabolism of specific amino acids. The microorganism has an absolute growth requirement for iron, preferentially in the form of haeme or its Fe(III) oxidation product haemin and when grown under conditions of excess haemin is highly virulent in experimental animals. A number of virulence factors have been implicated in the pathogenicity of *P. gingivalis* including the capsule, adhesins, cytotoxins and extracellular hydrolytic enzymes.

In order to develop an efficacious and safe vaccine to prevent *P. gingivalis* colonisation it is necessary to identify and produce antigens that are involved in virulence that have utility as immunogens to generate neutralising antibodies. Whilst it is possible to attempt to isolate antigens directly from cultures of *P. gingivalis* this is often difficult. For example as mentioned above, *P. gingivalis* is a strict anaerobe and can be difficult to isolate and grow. It is also known that, for a number of organisms, when cultured *in vitro* that many virulence genes are down regulated and the encoded proteins are no longer expressed. If conventional chemistry techniques were applied to purify vaccine candidates potentially important (protective) molecules may not be identified. With DNA sequencing, as the gene is present (but not transcribed) even when the organism is grown *in vitro* it can be identified, cloned and produced as a recombinant DNA protein. Similarly, a protective antigen or therapeutic target may be transiently expressed by the organism *in vitro* or produced in low levels making the identification of these molecules extremely difficult by conventional methods.

With serological identification of therapeutic targets one is limited to those responses which are detectable using standard methods such as Western Blotting or ELISA. The limitation here is the both the level of response that is generated by the animal or human and determining whether this response is protective, damaging or irrelevant. No such limitation is

present with a sequencing approach to the identification of potential therapeutic or prophylactic targets.

It is also well known that *P. gingivalis* produces a range of broadly active proteases (University of Melbourne International Patent Application No PCT /AU 96/00673, US Patent Nos 5,475,097 and 5,523,390), which make the identification of intact proteins difficult because of their degradation by these proteases.

SUMMARY OF THE INVENTION

10

The present inventors have attempted to isolate *P. gingivalis* nucleotide sequences which can be used for recombinant production of *P. gingivalis* polypeptides and to develop nucleotide probes specific for *P. gingivalis*. The DNA sequences listed below have been selected from a large number of *P. gingivalis* sequences according to their indicative potential as vaccine candidates. This intuitive step involved comparison of the deduced protein sequence from the *P. gingivalis* DNA sequences to the known protein sequence databases. Some of the characteristics used to select useful vaccine candidates include; the expected cellular location, such as outer membrane proteins or secreted proteins, particular functional activities of similar proteins such as those with an enzymatic or proteolytic activity, proteins involved in essential metabolic pathways that when inactivated or blocked may be deleterious or lethal to the organism, proteins that might be expected to play a role in the pathogenesis of the organism eg. red cell lysis, cell agglutination or cell receptors and proteins which are paralogues to proteins with proven vaccine efficacy. DNA sequences that were considered to be poor vaccine candidates and not selected include those that code for proteins involved in replication, non-essential proteins involved in cellular processes and those proteins present at sites that would be unlikely to be affected by

15
20
25

immune mediators such as those found in the bacterial cytoplasm or inner membranes.

In a first aspect the present invention consists in an isolated *P. gingivalis* nucleotide sequence, the nucleotide sequence consisting of or including a sequence selected from the group consisting of SEQ ID NO: 1 to 11 fragments thereof and sequences complementary thereto.

In a second aspect the present invention consists in an isolated *P. gingivalis* polypeptide, the polypeptide being at least partially encoded by a nucleotide consisting of or including a sequence selected from the group consisting of SEQ ID NO: 1 to 11, fragments thereof and sequences complementary thereto.

In a third aspect the present invention consists in an isolated *P. gingivalis* polypeptide consisting of or including an amino acid sequence selected from the group consisting of SEQ ID NO: 12 to 33 or fragments thereof.

In a fourth aspect the present invention consists in an isolated polynucleotide, the polynucleotide encoding a polypeptide consisting of or including an amino acid sequence selected from the group consisting of SEQ ID NO: 12 to 33 or fragments thereof.

In a fourth aspect the present invention consists in a nucleotide probe specific for *P. gingivalis*, the probe including a detectable label and a nucleotide sequence of at least 15 nucleotides, the nucleotide sequence being derived from a sequence selected from the group consisting SEQ ID NO: 1 to 11 and sequences complementary thereto.

In a fifth aspect the present invention consists in a composition for use in raising an immune response in an animal directed against *P. gingivalis* the composition including an acceptable carrier and/or adjuvant and at least one polypeptide having a sequence selected from the group consisting of SEQ ID NO: 12 to 33 and fragments thereof.

As will be understood by those skilled in the art the nucleotides of the present invention may be useful in DNA vaccination to reduce the incidence and/or severity of *P. gingivalis* infection.

Accordingly in a sixth aspect the present invention consists in a
5 composition for use in inducing an immune response, the composition including at least one DNA molecule, the at least one DNA molecule having or including a sequence selected from the group consisting of SEQ ID NO: 1 to 11 fragments thereof and sequences complementary thereto.

Further information regarding DNA vaccination may be found in
10 Donnelly *et al*, Journal of Immunological Methods 176(1994) 145-152, the disclosure of which is incorporated herein by reference.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group
15 of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

DETAILED DESCRIPTION

20 Preparation of the *P. gingivalis* library for sequencing.

To determine the DNA sequence of *P. gingivalis* genomic DNA was isolated from *P. gingivalis* strain W50 (ATCC 53978) essentially by the method described by Mamur J. (1961)(1). Cloning of DNA fragments was performed
25 essentially as described by Fleischmann *et al.*, (1995)(2). Briefly, purified genomic DNA from *P. gingivalis* was nebulized to fragment the DNA and was treated with Bal31 nuclease to create blunt ends then run twice on preparative 1% agarose gels. DNA fragments of 1.6-2.0 kb were excised from the gel and the DNA recovered. This DNA was then ligated to the vector
30 pUC18 (*Sma*I digested and dephosphorylated; Pharmacia) and

electrophoresed on a 1% agarose preparative gel. The fragment comprising linear vector plus one insert was excised, purified and this process repeated to reduce any vector without insert contamination. The recovered vector plus insert DNA was blunt-ended with T4 DNA polymerase, then a final
 5 ligation to produce circular DNA was performed. Aliquots of Epicurian Coli Electroporation-Competent Cells (Stratagene) were transformed with the library DNA and plated out on SOB agar antibiotic diffusion plates containing X-gal and incubated at 37°C overnight. Colonies with inserts appeared white and those without inserts (vector alone) appeared blue.
 10 Plates were stored at 4°C until the white clones were picked and expanded for the extraction of plasmid DNA for sequencing.

DNA sequencing

Plasmid DNA was prepared by picking bacterial colonies into 1.5ml of
 15 LB, TB or SOB broth supplemented with 50-100ug/ml Ampicillin in 96 deep well plates. Plasmid DNA was isolated using the QIAprep Spin or QIAprep 96 Turbo miniprep kits (QIAGEN GmbH, Germany). DNA was eluted into a 96 well gridded array and stored at -20C.

Sequencing reactions were performed using ABI PRISM Dye
 20 Terminator and ABI PRISM BIGDye Terminator Cycle Sequencing Ready Reaction kits with AmpliTaq DNA polymerase FS (PE Applied Biosystems, Foster City, CA) using the M13 Universal forward and reverse sequencing primers. Sequence reactions were conducted on either a Perkin-Elmer GeneAmp 9700 (PE Applied Biosystems) or Hybaid PCR Express (Hybaid,
 25 UK) thermal cyclers. Sequencing reactions were analysed on ABI PRISM 377 DNA sequencers (PE Applied Biosystems).

The sequences obtained are set out below.

SeqID#1-11 represents the DNA sequence that encodes proteins in SeqID#12-33. Proteins in SeqID#12-22 are the entire open reading frame
 30 from DNA SeqID#1-11. Proteins in SeqID#23-33 are the proteins encoded by

DNA SeqID#1-11 from their putative initiation codon. The initiation codon was calculated from sequence homology alignment using FastX or by the ORF prediction program GeneMark.

As will be understood by those skilled in the art open reading frames (ORFs) may be readily identified. ORFs may be determined using two methods, for example, alignments from FastX search results may be used to define the start and end positions of coding regions if sufficient protein homology is present. Alternatively, protein coding regions may be identified using the ORF identification program GeneMark (3) using a training matrix based on published *P.gingivalis* sequences. This matrix may be further refined by adding ORFs identified from the results of homology searching and ORFs identified by GeneMark. The program PSORT (4) may be used for the detection of signal sequences and the prediction of protein cell localisation. A UNIX version of TopPred (5) may be used to identify potential membrane spanning domains.

DNA sequence analysis

Raw trace data files from the ABI 377 sequencer were manually trimmed using Staden Pregap (Laboratory of Molecular Biology, Medical Research Council, UK) running on a Sun Microsystem computer. Trimmed files were assembled into contigs using Staden Gap v4.1 and exported as a consensus file in FastA format. This consensus was converted into GCG format files and analysed for homology using the FASTX algorithm on a non-redundant protein database compiled by ANGIS (Australian Genomic Information Service, University of Sydney). Individual FASTX search results were examined for significant homology by statistical probability and amino acid alignments.

The results are set out in Table 1.

It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to
5 be considered in all respects as illustrative and not restrictive.

Dated this twenty-ninth day of July 1998

CSL LIMITED

Patent Attorneys for the Applicant:

F B RICE & CO

References.

1. Mamur, J. (1961) A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J. Mol. Biol.* 3, 208-218.
5
2. Fleishmann, R.D. et al. (1995) Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496-512.
3. Borodovsky M, Rudd KE, and EV Koonin. (1994). Intrinsic and
10 extrinsic approaches for detecting genes in a bacterial genome. *Nucleic Acids Res.* 22:4756-4767.
4. Horton, P. and Nakai, K. (1996). A probabilistic classification system for predicting the cellular localization sites of proteins. *Intellig. Syst. Mol.*
15 *Biol.* 4: 109-115.
5. Claros MG and G von Heijne. (1994). TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* 10: 685-686.
20

Table 1

SeqID#	Length of protein in SeqID	Homology description	Length of protein homolog	% identity	Overlap	E value
1,12,23,34	843aa	Cysteine protease/hemagglutinin, Porphyromonas gingivalis	886aa	35%	835aa	3.20E-104
2,13,24,35	1175aa	Internalin, Lysteria monocytogenes	334-821aa (1)	39%	326aa	2.90E-30
3,14,25,36	312aa	TonB-linked adhesin, Porphyromonas gingivalis	1097aa	46%	351aa	2.30E-47
4,15,26,37	757aa	Ferric receptor, Campylobacter coli	696aa	23%	532aa	1.50E-15
5,16,27,38	569aa	Protease, Helicobacter pylori	459aa	35%	357aa	2.70E-45
6,17,28,39	290aa	Fimbrial motif (2)	N/A	N/A	N/A	N/A
7,18,29,40	1017aa	Outer membrane protein, Bacteroides thetaiotaomicron	1038aa	27%	1087aa	1.40E-35
8,19,30,41	811aa	Clp protease subunit, Bacillus subtilis	810aa	53%	662aa	6.70E-147
9,20,31,42	293aa	Cysteine protease, Porphyromonas gingivalis	1358aa	46%	114aa	8.20E-17

Table 1 (cont)

SeqID#	Length of protein in SeqID	Homology description	Length of protein homolog	% identity	Overlap	E value
10,21,32,43	419aa	Hemolysin, Helicobacter pylori	449aa	30%	440aa	1.40E-33
11,22,33,44	853aa	Surface antigen gene, Methanosarcina mazei	783aa	40%	281aa	1.40E-37

1 Int malin is a family of proteins in Lysteria monocytogenes of various sizes containing common repetitive motifs.
2 A small fimbrial motif was identified by amino acid sequencing. The entire protein contains no significant homology with other fimbrial prot ins.

(2) INFORMATION FOR SEQ ID NO:1

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2547 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Porphyromonas gingivalis*

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...2547

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1

25 TTCGGAATAT CACCTTCAAT GAAAAAAGT TTTCTTTTAG CCATAGTAAT GCTCTTTGGC 60
 ATTGCCATGC AGGGACATTC TGCTCCGGTT ACGAAAGAGC GAGCTTTGAG TCTGGCTCGG 120
 CTGGCTTTGC GACAGGTATC CTTGCGAATG GGACAAACAG CAGTATCTGA CAAGATTTCC 180
 ATCGATTACG TTTATCGGCA AGGAGATGCT GAGAGGGGTA TCACATCACA AGAGGAAGGC 240
 30 TCTCCTGCAT ATTTTTATGT AGCTAATCGT GGAAATAATG AGGGCTATGC TCTTGTAGCA 300
 GCAGATGACA GAATACCGAC AATTTTAGCC TATTCACCCA TTGGCCGTTT CGACATGGAC 360
 AGTATGCCGG ACAATCTTCG CATGTGGCTA CAAATTTACG ATCAGGAAAT AGGCCTGATA 420
 CTTTCCGGAA AAGCTCAGCT CAATGAAGAG ATATTACGTA CCGAGGGCGT ACCGGCTGAA 480
 GTACATGCTC TGATGGATAA CGGTCATTTT GCCAACGATC CCATGCGATG GAATCAAGGT 540
 35 TACCCATGGA ACAATAAGGA ACCACTGCTT CCTAATGGCA ATCATGCCTA TACCGGCTGT 600
 GTTGCTACTG CTGCAGCACA AATCATGCGC TACCATAGCT GGCCGCTTCA AGGTGAAGGC 660
 TCTTTCGATT ATCATGCAGG TTCATTAGTT GGCAACTGGT CCGGCACATT TGGTGAAGTG 720
 TACGACTGGA TCAATATGCC CGGAAATCCC GACCTTGATA ATCTGACTCA ATCTCAAGTG 780
 GATGCCTACG CCACACTGAT GCGTGATGTG AGTGCATCTG TTTTCGATGAG TTTTATGAA 840
 40 AATGGAAGTG GTACGTACAG CGTTTATGTA GTAGGAGCCT TGCGAAACAA CTTTCGCTAC 900

	AAGCGTTCAC	TGCAGCTACA	TGTACGCGCC	TTATATACCT	CACAGGAGTG	GCACGATATG	960
	ATCCGCGGGG	AACTTGCCTC	CGGAAGGCCG	GTCTATTATG	CAGGGAATAA	CCAGAGCATA	1020
	GGACATGCTT	TCGTTTGC GA	TGGTTATGCT	TCGGATGGTA	CTTTCCATTT	CAACTGGGGT	1080
	TGGGGAGGTG	TTTCCAACGG	CTTCTACAAA	CTAACACTCC	TCTCGCCGAC	TTCGTTGGGT	1140
5	ATCGGAGGTG	AGGGAATAGG	TTTTACCATT	TATCAAGAGA	TCATCACCGG	TATCGAACCG	1200
	GCTAAGACTC	CCGCTGAAGC	CGGTACAGAT	GCCTTGCCGA	TCTTGCCACT	GAAAGACATA	1260
	GAAGCCGAGT	ATAAAAGTGA	ATCCGGATTG	AACGTAGGGT	ATTTCGATATA	TAATACAGGT	1320
	GAAGAGCAAT	CAAATCTTGA	CCTCGGATAC	AGATTGAACA	AGGCTGACGG	AGAAGTCATA	1380
	GAGGTGAAAA	CTTCATCTAT	CAATATCTCT	TGGTACGGAT	ACGGAGAGCA	TCCCGAGAGT	1440
10	TTCTCATTGG	CACCTAATCA	GTTGTCACAA	GGAATCAACA	CCATCACCCCT	ACTTTATCGT	1500
	CGCACAGGCA	CCGAACAGTG	GGAGCCGGTA	CGGCATGCAC	AGGGAGGATA	TGTCAATAGC	1560
	ATTAAAGTAA	ATACGACAGA	CCCGAACAAAT	GTCGTAGTCA	CGGTAGATAA	TAACGAAGGC	1620
	AAGCTCAGTA	TCGTCCCCAA	CAGCTTTGTC	GCAGATCTGA	ATTCTTATGA	ACATAGTACG	1680
	ATTACAGTAC	AGTTCAATAG	CGACAGCCCT	GATGAGATCC	GTACACCCGT	AGCCTTTGCT	1740
15	CTATCTACAG	GAGCTACTGC	GGACGATGTA	ATATCTTTGG	GCTGGGTAAT	GGCTGAAGTT	1800
	CCGGGCGGTA	GCAGCAACTA	TCCGGTGGTT	TGGTCTAAAG	ACGTTCTCAC	TCTCTCGGAA	1860
	GGCGACTATA	CATTGTGGTA	TAGATTTTCC	ATCAACAACC	AAAAGGATGA	ATGGAAAAAG	1920
	ATCGGAAGCG	TGTCAGTAAA	AACACCGACA	GAGTATACGC	ACCCCTTATT	CGAAGTGGGC	1980
	CATAATCAAA	CTTCTACCTA	TACGCTGGAT	ATGGCACACA	ACAGAGTATT	GCCCGACTTT	2040
20	AACTCAAAA	ATCTCGGATT	GCCTTTCAAT	GGTGAGTTGG	TTGTTGTTTT	CCGCCAAACA	2100
	CAATCCTCAT	CGGGGTCTTT	ATGGGCAGCT	CAAGAAACAG	TACATATCAA	GCAAGGAGAA	2160
	ACTTTCGTAT	ATAAACCTGT	TGTCGAAGGC	CCTATACCTG	ATGGATCCTA	TCGTGCGACC	2220
	CTCCATGCAT	TCGTAAACGG	ACAACAACAG	TTGTACCTCA	AGGGGAAAAG	GAAC TACACG	2280
	GTGAAGATCG	TCAATGGTAC	AGCGGTAGAA	GCAATAGAAT	CGTCAGAAGA	GATCAGAGTA	2340
25	TTCCCTAATC	CGGCACGCGA	TTATGTGGAA	ATATCGGCAC	CTTGCAATTCC	CCAAGAAACA	2400
	TCTATCATTC	TTTTTCGATCT	GTCAGGCAAG	ATTGTCATGA	AGAATAGTTT	ATCAGCGGGG	2460
	CATGGCAGAA	TGGATGTCAG	CCGACTTCCT	AATGGGGCCT	ACATCCTTAA	GGTGGATGGA	2520
	TATACGACGA	AAATAAATAT	AGTGCAC				2547

30

(2) INFORMATION FOR SEQ ID NO:2

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 3807 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

5 (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

10 (B) LOCATION 1...3807

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2

	GTAAATGTA TGGGCAAGTA TAAAAGAGCT AAGTACCGCT ATTGGCTTTT TCCTTTCTGT	60
15	TCGGATTATT ATACCTTTGA GGGAGTTACT TTTTATGCG CATCTGACGA TATGACAACC	120
	AAGAAACCCC AAGCCATTTT AGACTTAGAG AAGGCCTATA ACATTGAAAT TCCTGATCTC	180
	TCCTCACAAG AAGGGATAAG CTGGTCGGTA AATAGATATT TCAAGCAAGA TTCTCCGGT	240
	GCAGTCGTTG AGCTTTGCTT GCGAGAATGT CAGATAGAAA GCATGACTTG GCTTATTGAT	300
	TTTCCTGCTC TAAAAAGCT TGATCTATCG TATAACCAA TCAGTAAGCT AGAGGGTCTA	360
20	GAACGTCTTA CTTCGTTAAC AAAACTTCGT CTAAGAAGTA ACCAAATCCG TAACTAGAG	420
	GGCCTGGATA GTCTCACCTC GCTAACAAA CTTTCTCTCT CCGATAACCA AATCAGTAAG	480
	CTAGAGGGTC TGGAACGTCT CACCTCGTTA GCGGAGCTTT ATCTTTTGGA TAACCAAATC	540
	AGTAACTAG AGGGTCTGGA ACGTCTCACG TCCTTAGCAA CGCTTGAAC ATCGGGTAAC	600
	CAAATCCGTA AGCTGGAGGG TCTGGAACGT CTCACGTCCT TAGCAACGCT TGAACATATCG	660
25	GGTAACCAA TCCGTAAGCT AGAGGGTCTG GAACGTCTCA CTTCGTTAAC AAAGCTTCGT	720
	CTAAGAAGTA ACCAAATCAG TAAGCTAGAG GGTCTGGAAC GTCTCACGTC CTTAGCAACG	780
	CTTGAACTAT CGGGTAACCA AATCCGTAAG CTGGAGGGTC TGGAACGTCT CACGTCCTTA	840
	GCAACGCTTG AACTGTCGGG TAACCAAATC AGTAAGCTAG AGGGTCTGGA ACGTCTCTCT	900
	TCGTTAACAA AGCTTCGTCT AAGAAGTAAC CAGATCAGTA AACTAGAGGG CCTGGAACGT	960
30	CTCACCTCGC TAACAAAAC TTTCTCTCTCC GATAACCAA TCAGTAAGCT AGAGGGTCTG	1020
	GAACGTCTCA CCTCGTTAGC GGAGCTTTAT CTTTGGATA ACCAAATCCG TAAGCTGGAG	1080
	GGCCTGGAAC GTCTCACCTC GTTAACAAAG CTTGCTCTAA GAAGTAACCA AATCAGTAAA	1140
	CTAGAGGGCC TGGATAGTCT CACCTCGCTA ACAAACCTT CTCTCTCCGA TAACCAAATC	1200
	AGTAACTAG AGGGCCTGGA ACGTCTCACG TCCTTAGCGG AGCTTTATCT TTTGGATAAC	1260
35	CAAATCCGTA AGCTGGAGGG TCTTGATGGT CTGCTTCCT TAACAAGGCT TAGTCTAAGG	1320
	CGCAACCAA TCAGTAAGCT GGAAGGACTA GACAGACTAA AGGTTTTGAG AAAACTTGAT	1380
	GTTTCGGGCA ATGATATTCA ATCTATTGAT GATATTAAGC TATTGGCTCC GATTCTGGAG	1440
	CAAACCTTAG AAAAAGTGA AATCCATGAC AATCCATTTG TTGCATCATC AGGCTTGATA	1500
	CTCTCTCCTT ATGATAATCA TTTGCCGGAG ATTAAAGCTC TTCTTGAAAA AGAAAAAGAA	1560
40	AAACAGAAAA AGACTTCAGT TGAATATCAC CCATTTTGCA AAGTAATGCT ATTGGGAAAT	1620

	CATTCTTCGG	GTAAAACAAC	ATTTCTTAGT	CAATACGATA	CAAATTATAC	GTATCAGAAA	1680
	AATACACATG	TGTTGTCGAT	ACATCGAAGC	AATAACCCTA	ATGCGATCTT	TTACGACTTT	1740
	GGGGGACAGG	ACTATTATCA	TGGGATTTAC	CAAGCCTTTT	TTACCACCCA	ATCGTTATAC	1800
	CTTCTCTTTT	GGGATGCTAA	GAAGGATCGA	AACTTTGTGA	GCGTAGATGA	TAAAGAATAT	1860
5	CAGACTCTTA	ATTTCAATCG	CCCCTATTGG	TTAGGACAGA	TAGCCTATGC	CTGCAATCGT	1920
	TGTATGTCCG	TTGGAGGAAA	TCCTGATGGC	AAGGACACAC	CACAGACCAC	AGACGATACA	1980
	ATTATCATTC	AGACTCATGC	CGATGAAACG	GGCGCTAAGC	AGCAAACCTT	AGGCTGTGCA	2040
	GCCGAGAATG	GAGTATTGGA	AGAAATCTAT	GTATCCTTAG	AGCCCAAGGC	GAATAGTGCC	2100
	GTACATGCGC	TCAACTATCT	GAATGAGCGG	GTGCGAGAAG	TTGTCGCAAG	CAGGAGTAAA	2160
10	TCAATTCAGA	TCACAGAAAA	AGATAAGGGA	TTGTACGAAG	CTCTTCCCAC	AATCGCCGGT	2220
	GATAATAAAC	ACATCCCTAT	CTCTCTCGAA	GCTCTTGCGG	CTCAATTGAA	TAAGGGAAGA	2280
	GCTGAAAATG	ATCTTTACAC	CATAGAGTAT	CTACAGACCG	AATTGAACCA	GCTTAGTCTG	2340
	CGAGGGGAGG	TGCTTTACTA	TCGTGAGAAT	GAGAAGCTGA	ACAAATTATGT	CTGGTTAGAT	2400
	CCGGCAGCTT	TTGTCCAAAT	GATTCATGGA	GAAATCCTCC	AAAAAGACAA	CATCAATAGA	2460
15	GGAACAGTTC	CTAAAGACAT	TTTTGAATGC	AAACTGCATA	ATCTAAGTTC	CGGAAGTATA	2520
	TTTGAAGAAG	ATGGCCAAAA	TGGTAATATG	ATCTTGCAGC	TATTATTGGA	AGAGCTGATC	2580
	GTATATGAAG	ATAAGGACTG	CTATGTGATA	CCGGGCTATC	TCCCTTTGCA	TTCCGATGAC	2640
	GAAGCCTATA	AATGGCTTAC	TTTGGGATTC	GAGAGGCCCA	ATTTTGTCTT	CAAATTCGAA	2700
	CGTTTTATCC	CCTTTGGCCT	GATCAACCAG	ATTATAGCCT	ACTATGGCCG	GGAAGAAGGT	2760
20	GCTCTAAAGC	GGTATTGGCG	AGATCAGGTC	ATCTTCACAG	CAGGCCGTGA	GATGGATAGG	2820
	CAAACGCTTG	AGCAAGAAGA	AGAGAAAGAG	GGTTTGCCCA	AGACGAATGC	CGAGGATTAT	2880
	CAGATCTGGA	TCAAGCTCGA	CTTTACCGAC	TTGGCCATAT	CCGTATTTCAT	CAAAGAGCAG	2940
	AGAAAGACAT	CAGCTAAGGA	TATGCAGCGG	AAAGAGGCTA	CTATCCTCAG	TGATATGTTG	3000
	GATATGTATT	GGAACAATAT	CCCTCCGAGG	GAGCAAATAG	GAGATAAGGA	TACGGAGCAA	3060
25	ACGAGAAGCA	CTATTCGTGA	AACAAACAGA	AAGAAGAGAC	CCATCCAGGA	TCTCTACCTC	3120
	TCCTGTGCCC	AAGCGGATAA	AGATTTGACG	GAGTCTCATT	ATATCCATTT	GGGCACGCTG	3180
	GACGATGAAA	GCAAGACTAC	GGCGAGGATT	GCAGCCTATC	CGTTGAAGAA	CGGCGTTATC	3240
	GATAAAGAGC	GGGTGCGAGA	AGTATCGACT	CGTCCCTACA	AACATCTTTC	CGTCAATAAA	3300
	AATCTGGCTA	CTGCAAAACA	GATCTTTATT	TCCTATTCCA	AAGAGGATCA	GACTGAACCTG	3360
30	GAGACCTGTC	TGCAATTTTT	CAAACCCTTG	GAGAAGAATG	GTCAGATCGA	GATCTACTAT	3420
	GATAAGTTGA	CTAAGTTTGA	AACACCTATT	CACCCTGAAA	TAAGAAAGCG	TATTGTGCGAA	3480
	GCCGACTGTA	TAATCGCTTT	GATCAGCCAA	CGCTATCTGG	CCACGGATTA	CATCCTGGAT	3540
	CATGAGTTGC	CTGTATTTCG	GGAGTATAAC	AAGACCATAG	TGCCGATATT	GATCAAGCCT	3600
	TGTACATTCG	AAGACGATGA	GTTCCCTTCGG	GAGAAAATATT	TTGCTCAGAA	AGCTCAAATA	3660
35	ATCAATCTTG	GAAAAGAGGG	AAAAACCATT	AAAGCTTATG	ATAGTATTAC	GGCATCAGCC	3720
	CATCGTGATG	AAAATTGGGT	GGCAGTAGTC	AGAGAGTTCA	AAGAGAAGAT	ATTAAGAATA	3780
	ACAAAACAGG	AGGTAAATAC	AGATGAA				3807

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1131 base pairs
 - (B) TYPE: nucleic acid
 - 5 (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- 10 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
- 15 (A) ORGANISM: *Porphyromonas gingivalis*

- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...1131
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3

	TTGTTTAATC ATAAAAATC ATGGTATGAA ACATTTCAAT TTTATCTCGT TGTTTTCCGC	60
	TCTGGCTTTA TTCTTTTGTG TGAAAATACC CTTGCACAAC AAAAAACAGA GGAGTTTGCA	120
25	CCTGTGTCGG ATTTACGTGC AGAAGCGTAC GGCTCTACCG TTTTCCTCCA CTGGACTCCG	180
	CCGTATGACA ATCCGATGAT TCCTCTAAGC GAGAGTTTTG AATCAGGTAT TCCAGCTATA	240
	TGGAAGACCA TTGACGCAGA TGGCGATGGC TATAATTGGA TGCATTTGAC CAATTTACAG	300
	GGACAGAGTG GTCTCTGTGT CTCTTCGGCT TCATACATAG GCGGCGTCGG AGCTTTGACT	360
	CCGGACAATT ATCTGATAAC ACCCGAATTA AACTACCCA CAGACGCGTT GGTGGAAATA	420
30	ATCTATTGGG TATGTACTCA AGATCTCACT GCTCCATCGG AGCACTATGC CGTTTATTCC	480
	TCTTCTACAG GCAATAATGC TGCTGACTTT GTTAATCTCT TATATGAAGA GACTTTGACT	540
	GCCAAACGGA TACAATCCCC CGAGTTGATC CGCGGAAATC GGACACAAGG TGTTTGGTAT	600
	CAAAGAAAGG TGGTACTCCC TAACGATACT AAATATGTTG CTTTCCGCCA TTTTAATTCC	660
	ACGGATAATT TCTGGCTCAA TTTGGATGAA GTATCTATCC TGTATACCCC TCTTCCCCGA	720
35	AGAGCTCCGT GTCCGCATCC GGGTGGTTAC ACTTATTCTG TATTCCGTGA TGGACAAAAG	780
	ATAGCGAGTG GATTGTCGGC ATTGGCATAT ATCGATACGG ATGTACCGTA TGGGACTCAA	840
	GACTATTGTG TCCAAGTCAA TTATCTGCAA GGAGACTCGT ATAAAGTCTG CAAAAATATA	900
	GTGGTGGCAA ATTCTGCAAA CATCTATGGG GCGGATAAGC CTTTTGCGTT GACCGTGGTT	960
	GGCAAGACCA TTGTAGCGAG TGCTTTCAAA GGAGAGATCA CTCTTTATGA CATTCGTGGC	1020
40	CGGCTGATAG CTTCCGGCTG CGATACGCTT AGGTACAAAG CGGAAAATGG TTTTACCTC	1080

ATTAAAATAC AGGTAAACGG AACTGTCTAT ACTGAGAAAA TCCAAATCCA A

1131

(2) INFORMATION FOR SEQ ID NO:4

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2289 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

10

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

20

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...2289

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4

	CATTTTTTAG TAACGATTAT GCGGACAAAA ACTATCTTTT TTGCGATTAT CTCTTTTATT	60
	GCTCTATTGT CGTCTTCTCT GTCGGCTCAG AGCAAAGCCG TTTTAACCGG TAGTGTGTCG	120
	GATGCCGAAA CCGGAGAGCC TCTTGCCGGT GCTCGAATCG AAGTCAAACA CACCAACATA	180
30	GTAGCCGGTG CCGATGCCGG CGGACATTTT GAGATCAAGA ACCTGCCGGC AGGGCAGCAT	240
	ACTATTATAT GTTCGTTGGG GGGGTATGGA CAGAAAGAGG AGGTGGTTGC CATCGAAGCC	300
	GGACAGACCA AAACGATCTC TTTTGCAATTG CGACTGCGAA CGAACAACTT GGAGGAAGTC	360
	GTCGTTACCG GTACCGGTAC ACGTTACCGC TTGGTCGATG CTCCTGTGGC AACGGAAGTC	420
	CTTACCGCTA AGGACATAGC CTCTTTCTCG GTCCTACTT CCGAGGCCTT ATTGCAGGGG	480
35	CTGAGTCCGT CTTTGGACTT CGGCCCAAT CTGATGGGCT CTTTCATGCA GCTGAACGGC	540
	CTTAGCAGTA AGTATATCCT CATCCTTATC GATGGTAAGC GTGTGTACGG CGATGTAGGC	600
	GGTCAGGCCG ATTTGAGTCG TATTTCTCCT GATCAGATCG AACGGATCGA ACTGGTGAAA	660
	GGTGCTTCGA GTTCGCTCTA CGGATCCGAT GCCATCGCCG GGGTAATCAA TGTGATCACA	720
	AAAAAGAATA CGAATCGACT GAGTGCATAT ACGTCACATC GCATATCGAA GTACAACGAT	780
40	CGGCAAACCA ATAATTCTGCT CGATATAAAC ATCGGTAAGT TCAGTAGCAA TACCAACTAT	840

	TTCTTCTACC	ATACGGATGG	CTGGCAGAAT	AGTCCGTTCCG	AAATAAAAAA	GAAAAAAGGA	900
	TCCGGCGAAC	CGGTCTTGGA	GGAAACGTAT	AAGAAAACCTT	TTCGTGCACA	GGAAAATCAG	960
	GGTGTAAAGCC	AATCGCTTTC	CTATTATGCA	ACTAACAATC	TTAGCTTCAG	CGGAAATGTG	1020
	CAGTACAATA	AACGTCAGAT	CTTCACTCCG	ACTTTTTTCCG	AAAAGAAGGC	CTATGACATG	1080
5	GATTATCGTG	CTTTGACGGC	TTCACCTCGGT	ACGAACATATC	TTTTCCCCAA	TGGTCTGCAT	1140
	ACGCCTTTCTT	TCGATGCCGT	CTACGATCGC	TTCCGTTTCCG	GATATTTGTA	TCATGACAAG	1200
	GACAGCAGTG	AGAGCCTGAT	CAACAACCAA	GGTCAGACCG	AGCAACCCAC	ATTCTTTCCG	1260
	GGTCAGCTAC	GCAATAAAAA	CGATCAGATC	CGATACACGG	CAGAGGCTCG	CGGTGTATTT	1320
	AACTGTCCTT	ATGCGCAGAA	ACTGACCGGC	GGTTTGGAGT	ATTTCCGTGA	GGAATTGATC	1380
10	TCTCCCTATA	ATTTGATTAC	CGACAAGGCA	GATGCTTCCA	CGCTCTCTGC	TTATGTACAA	1440
	GATGAATGGA	AACCGCTCGA	TTGGTTCAAT	ATGACAGCCG	GTTTCCGTCT	GGTACACCAT	1500
	CAGGAGTTTCG	GTACACGAAT	GACGCCTAAG	GTATCCATAC	TCGCCAAGTA	TGGGCCCGCTG	1560
	AACTTCCGCG	CTACGTATGC	TAACGGCTAT	AAGACTCCCA	CGCTGAAAGA	GCTTTTTTGCA	1620
	CGGAACGAAC	TCACCACTAT	GGGTTTCGCAC	AATCTCTATC	TCGGCAATGC	GGATCTTAAG	1680
15	CCACAGATGT	CGGATTATTA	TGCTTTGGGC	TTGGAGTACA	ATCAAGGCCC	TATCTCGTTC	1740
	AGTGCAACGG	TTTATGACAA	TGAACTTCGC	AATCTGATCT	CCTTTATGGA	TATACCGACC	1800
	TCACCCGAGC	ACGAAGCTCA	GGAATCAAG	AAAACCAAGC	AGTATGCCAA	CATAGGAAAA	1860
	GCTCGCAGCC	GCGGCCTTGA	TGTCCTATGT	GATGCCTCTA	TCGGTTGGGG	TATCAAGTTA	1920
	GGAGCCGGAT	ACAGCCTCGT	GGAAGCTAAG	AATCTCCAGA	CGGATGAGTG	GCTGGAAGGA	1980
20	GCTGCACGTC	ATCGTGCCAA	TGTGCACGCC	GATTGGGTTC	ACTACTGGGG	TCAGTATAGA	2040
	CTTGGCGTGA	GCCTTTTCGG	CCGTATTTCAG	AGCGAGCGTT	ACTACAAAGA	CGGCAATGCT	2100
	CCGGACTATA	CCTTGTGGCG	ACTCGCCACA	TCGCATCGTT	TCGCTCATTT	CCGCCACATC	2160
	ATCCTGGATG	GAACGCTCGG	TATAGACAAC	CTGTTTGACT	ACGTGGATGA	TCGTCCTATG	2220
	GGTGTCAATT	ATGCTACCGT	AACGCCGGGA	CGTACTTTCT	TTGCTCAAAT	AGCGATTCTGA	2280
25	TTCAACAAC						2289

(2) INFORMATION FOR SEQ ID NO:5

- 30 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1746 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: circular
- 35 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- 40 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

5 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1746

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5

10

	TCAAAAATAG	TACTACGAAA	GTTTTGTACC	TTAGCGCGCA	TGAAAAAGAC	CAATCTGTTT	60
	TTATCTCTGC	TGGTGATCTT	TATCACC GGT	AGTTTTATGA	CTGCCTGTGC	ACAGAAGTCC	120
	AAGACGAACA	AACTCACC GA	AGAAGATCGG	AGCCGCAATG	AGTATGTACA	GTCGATGGAT	180
	GTGCTTAGCA	ATATTATCGG	TAACGTCAGG	CTGTATTTCG	TCGATACCAT	AAGTATCAAA	240
15	CATATGACTC	GGCGTGGTAT	AGATGCGATG	TTGGGCGGGC	TTGACCCCTA	TACCGAATAC	300
	ATTCCCTTACG	AGGAAATGGA	TGAACTGAAA	TTGATGACTA	CGGGAGAGTA	TGCCGGAGTC	360
	GGAGCTATCA	TATCGCAGCG	CCCGGATAGT	GCTGTGATTA	TCCAGAGACC	TATGGAAGGT	420
	ATGCCCCGAG	ACGAAGCAGG	ATTGATAGCA	GGCGACCGCA	TCCTGACTAT	CGATGGGAAA	480
	GACTTCCGCA	AATCCACCAC	ACCGAAAGTA	AGCCAAGCAC	TGAAAGGGAT	AGCCGGTACT	540
20	GTTGCAAAGG	TGACAGTAAT	GCGCTATGGC	GAAACCAAAC	CTCGTACTTT	TTCCGTGAAA	600
	CGTCAAAAAG	TGATTATGAA	TTCCGCTCACT	TACAGCGGAA	TGCTCGATGG	CTCGATAGGA	660
	TATATCCGCT	TGAACAACCT	TACGGACAAA	AGTGCAGAAG	AGGTGCGCAC	GGCCTTGTTG	720
	GATCTTCGTG	ACAAACAAGG	AGCGAAAGGT	CTCATTTTGG	ATTTAAGAGG	CAATGGTGGC	780
	GGACTGATGC	AGGCTGCTAT	CGAGATAGTC	AATCTGTTTCG	TCCCTAAGGG	CAAAGAGGTG	840
25	GTAACGACCA	AAGGTCGCAT	TGCAGAGTCG	GCGTCCGTAT	TTCGCACATT	GACTGAACCG	900
	ATCGACACGA	AACTCCCGAT	AGTAGTCCTG	ATCGATGGAC	AATCGGCATC	TTCTTCGGAG	960
	ATTGTAGCCG	GAGCACTGCA	GGATATGGAC	AGGGCTGTAC	TGATGGGACA	AAAGAGCTAT	1020
	GGCAAAGGGC	TTGTACAAAC	GACTCGTCAG	CTACCATACA	ACGGCGTGAT	CAAATTGACT	1080
	ACGGCCAAGT	ACTACATCCC	AAGCGGACGT	TGTATTTCAGC	GTTTGGACTA	CAGCCGCACC	1140
30	AATCGGACAG	GTATGGCAAC	GGCCATTCCCT	GACAGTCTGC	ACAAAATCTT	TTACACTGCT	1200
	GCCGGAAGAC	GTGTAGAAGA	TGCAGGAGGA	ATCCTGCCTG	ACATCGAGGT	CAAACAAGAT	1260
	ACAGCTGCGA	CATTACTTTA	TTATATGGCC	ATCAATAATG	ACGTTTTTCGA	TTTCGTCACA	1320
	GGTTATGTGC	TCAAGCATAA	AACGATTGCC	AAGCCGGAGG	ATTTTTCCAT	AACGAACGAG	1380
	GACTATGCAG	CTTTCTGCAA	GATGATGGAA	GAAAAGAAAT	TTGACTATGA	TCGCCAGAGT	1440
35	GGCAAGATGC	TTGACAAACT	GGAGGAACTG	GCTAAGATAG	AAGGCTACCT	GCCGGAAGCC	1500
	AACTCGGAGC	TTAAAGCACT	ACGCGAAAAG	CTAAAACCCA	ACCTGTCGCG	TGATCTGCTA	1560
	CGATTCAAAA	AGGAGATAAC	AAACTATCTC	AACAATGAGA	TTGTCACTCG	CTATTATTAT	1620
	GAGCGAGGCA	GTATCCGCCA	GAGTTTGCCG	GAAGATAAGG	TAGTCAAAGA	AGCTATTAAG	1680
	CTGCTGAAGG	ACCATCCGGA	ACAAATTTCGA	CAGATCCTTG	CAGCTCCGAA	AGCAGAGAAT	1740
40	AAAGGG						1746

(2) INFORMATION FOR SEQ ID NO:6

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 885 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *Porphyromonas gingivalis*

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6

25	ACGAAGAAAA CACTGATGAA AAAGCTATTT CTCTCGCTCA CGAGTCTTGT AATGGTCTTC	60
	GCTGTTGCAA GTTGCGATAT AATCGACAAG GATCAAACCC TCTTGCCGGC TCCGACCAAT	120
	GTGACACCCG ATAATCCGGA TGACAATCCT TCGGAGATCG ACATTACGCA GACGCACACA	180
	GAAAAATATG TTTTGGCTGA AGAATTTACC GGCCAAAAAT GTCTCAACTG TCCGAAAGGT	240
30	CATCGCAAAC TGGCGGCTCT CAAGGAGCAA TACGGTAAGA GATTGACTGT TGTCGGTATA	300
	CATGCCGGCC CTGGATCTCT CGTGCCACCT CTTTTCCGTA CAGAAGCCGG AGACGCATAT	360
	TATAGCAAGT TCGCCAATAA TACCCCTCTC CCTGCGCTGA TGGTTTCGCG CAAAAAGTTC	420
	GGCTCTTCCT ACGTTTATGA TAAGAGCTAC AAAACGTGGG ACGTGCCTAT TGCCGAGCAG	480
	ATGGAGCAAA AGGCGAAGAT CAATATCTTT GCCGTGGCCG AATACACCGA TACCCAAAAG	540
35	ATCAAGGTGA CTGTAAAGGG TAAAATACTG GAGGGGAATA CACTCCCGAA GTCCATGGTT	600
	CAGGTGTATC TGTGGAGGA TAAGCTGATC GCTCCGCAGG TGGATGGCAA TACGACAGTC	660
	GAGAAATTACG AGCACAATCA CGTGTTGCGT GGAGCCGTTA ATGGTATTTG GGGCGAAGAA	720
	TTTGTGAATC TCAAAGATTA TTTGTATACT TACGCCGTTG AACCGCTCTC GGGTATGTCC	780
	TTCGTAGCCG AGAATTATTC GATTGTGGCT TTTGTATACG ATGTGCAGAC GTTCGAAGTG	840
40	TATGACGTTG TGCATGTAAA GATCAATCCG CAATCCGATG GCAAA	885

(2) INFORMATION FOR SEQ ID NO:7

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 3138 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

10 (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

15 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Porphyromonas gingivalis

20 (ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...3138

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7

25 GACTTCCCTT GGGTTAGGGT AAAACCGGAG AAGAAAAGAA AACAACATAA CAGTAATAAT 60
 TTTAAGTTTA ACGCAAAGA AAAGTCTATG AAAAGAATGA CGCTATTCTT CCTTTGCTTG 120
 CTGACGAGCA TTGGGTGGGC TATGGCCCAG AATAGAACCG TGAAGGGTAC AGTTATCTCC 180
 TCCGAGGATA ATGAGCCCCT GATCGGCGCG AATGTCGTGG TTGTCGGAAA CACCACTATC 240
 30 GGTGCTGCAA CCGACTTGGA TGGCAACTTC ACGCTTAGCG TGCCTGCCAA TGCCAAAATG 300
 TTGAGAGTGT CCTATTCCGG TATGACTACC AAAGAGGTCG CCATCGCTAA TGTGATGAAG 360
 ATCGTACTGG ATCCGGACTC TAAGGTCTG GAGCAGGTAG TTGTATTGGG TTACGGTACG 420
 GGACAGAAAC TCAGCACTGT TTCCGGTTCT GTGGCCAAAG TGTCCAGCGA AAAGCTCGCG 480
 GAAAAGCCCG TTGCCAATAT CATGGATGCC CTCCAAGGTC AGGTAGCCGG TATGCAGGTT 540
 35 ATGACTACAT CCGGTGACCC TACTGCCGTC GCTTCTGTGG AGATCCATGG TACAGGGTCG 600
 TTGGGGGCAA GCTCTGCACC ATTGTATATC GTGGATGGTA TGCAAACCTC TTTGGATGTT 660
 GTGGCTACGA TGAATCCGAA TGATTTTGAA TCTATGTCCG TTTTGAAAGA TGCTTCTGCA 720
 ACATCTATTT ATGGAGCTCG TGCTGCAAAC GGAGTCGTTT TCATTCAAAC GAAGAAAGGT 780
 AAAATGAGCG AGAGAGGTCG TATTACCTTT AATGCCAGTT ACGGGATTTT TCAAATCCTG 840
 40 AATACTAAGC CCCTTGATAA TATGATGACT GGAGATGAAT TGCTGGATTT TCAGGTGAAG 900

	GCAGGTTTTT	GGGGGAACAA	TCAAACCGTT	CAGAAAGGTTA	AAGATATGAT	CCTTGCCGGA	960
	GCTGAAGATT	TGTATGGCAA	TTATGATTCT	TTGAAAGATG	AGTATGGTAA	GACATTGTTC	1020
	CCAGTGGATT	TTAATCATGA	TGCAGACTGG	CTCAAGGCTT	TGTTTAAAC	AGCACCCACC	1080
	AGTCAAGGTG	ATATTTCTTT	CTCCGGAGGG	TCTCAGGGAA	CTTCATATTA	TGCCTCTATA	1140
5	GGCTACTTCG	ATCAGGAAGG	TATGGCTCGT	GAACCGGCAA	ATTTTAAGCG	CTATAGTGGC	1200
	CGGCTCAACT	TCGAAAGTCG	TATCAATGAA	TGGCTGAAAG	TTGGTGCAA	TTTGTCTGGT	1260
	GCGATAGCGA	ATAGACGATC	TGCCGACTAT	TTTGGAAGT	ATTATATGGG	GTCAGGTACT	1320
	TTCGGTGTGT	TAACGATGCC	TCGTTATTAT	AACCCTTTTG	ATGTGAATGG	GGATTTAGCA	1380
	GATGTCTATT	ACATGTATGG	AGCTACCAGA	CCTTCTATGA	CAGAACCGTA	CTTCGCAAAA	1440
10	ATGAGACCGT	TCAGTTCCGA	ATCACATCAG	GCCAATGTAA	ATGGTTTCGC	CCAGATTACT	1500
	CCGATCAAAG	GCCTTACTTT	AAAGGCACAG	GCTGGTGTTG	ATATTACTAA	TACTCGCACT	1560
	TCTTCTAAGA	GAATGCCCAA	TAATCCGTAT	GATTCTACTC	CTCTTGGGGA	AAGAAGAGAA	1620
	AGAGCTTATC	GAGATGTTAG	CAAGTCTTTT	ACAAATACGG	CTGAATATAA	GTTTTCAATT	1680
	GATGAAAAAC	ATGATCTTAC	AGCATTGATG	GGGCATGAAT	ATATTGAATA	TGAAGGGGAT	1740
15	GTTATTGGGG	CATCTTCTAA	AGGATTTGAA	AGTGATAAGT	TGATGTACT	GAGCCAGGGA	1800
	AAAACCGGAA	ATAGTTTGTC	TTTGCCTGAA	CACAGAGTCG	CTGAATATGC	CTATTTGTCT	1860
	TTCTTTAGTC	GTTTTAATTA	CGGTTTTGAC	AAATGGATGT	ATATAGATTT	CTCTGTTCGT	1920
	AATGACCAAT	CCTCTCGATT	CGGATCCAAT	AATAGAAGCG	CGTGGTCTA	TTCTGTCGGT	1980
	GGAATGTTTG	ACATATATAA	TAAATTCATT	CAAGAAAGTA	ATTGGCTCAG	TGATCTTCGA	2040
20	CTGAAAATGA	GTTATGGTAC	AACGGGTAAC	TCGGAGATTG	GTAATTACAA	CCACCAAGCA	2100
	CTCGTTACTG	TGAACAATTA	TACTGAAGAT	GCTATGGGGC	TTAGCATTTT	TACAGCAGGC	2160
	AATCCCGACC	TCTCGTGGGA	AAAGCAGTCT	CAGTTCAACT	TCGGTTTGGC	TGCAGGGGCT	2220
	TTCAATAATC	GCTTATCTGC	AGAGGTAGAT	TTCTATGTCC	GCACTACGAA	TGATATGTTG	2280
	ATTGATGTCC	CGATGCCTTA	TATCAGTGGT	TTCTTCTCAC	AGTATCAGAA	TGTAGGCTCT	2340
25	ATGAAAAATA	CGGGTGTAGA	CCTTTCTCTT	AAGGGGACGA	TCTACCAAAA	TAAGGACTGG	2400
	AATGTATATG	CTTCTGCGAA	TTTCAACTAC	AATAGACAGG	AAATAACAAA	GCTTTTCTTC	2460
	GGTCTCAATA	AGTACATGTT	GCCTAATACC	GGTACTATAT	GGGAAATTGG	GTACCCCAAT	2520
	TCGTTCTATA	TGGCTGAATA	TGCTGGAATC	GACAAAAAAA	CCGGTAAGCA	GTTGTGGTAT	2580
	GTTCTCTGGT	AAGTCGATGC	GGATGGTAAT	AAAGTTACAA	CAAGCCAGTA	CTCAGCTGAC	2640
30	TTGGAGACAC	GAATTGATAA	GTCTGTTACT	CCTCCTATTA	CAGGTGGTTT	CTCCTTAGGT	2700
	GCTTCTTGGA	AAGGACTTTC	TTTAGATGCT	GATTTTGCCT	ACATCGTTGG	TAAATGGATG	2760
	ATCAATAATG	ACCGTTACTT	TACAGAGAAT	GNAGGTGGAT	TGATGCAATT	AAATAAAGAT	2820
	AAAATGCTAT	TGAATGCCTG	GACAGAGGAT	AATAAAGAAA	CAGATGTTCC	AAAATTGGGA	2880
	CAGTCTCCTC	AGTTTGATAC	GCATTTGTTG	GAGAATGCTT	CTTTCCTGCG	TTTGAAGAAT	2940
35	CTCAAACTCA	CCTATGTACT	CCCCAATAGT	CTTTTTGCTG	GGCAGAATGT	GATTGGTGGA	3000
	GCTCGTGTCT	ATTTGATGGC	GCGCAATCTG	TTAACTGTTA	CGAAGTATAA	AGGCTTTGAC	3060
	CCTGAAGCAG	GGGGGAATGT	GGGAAAAAAT	CAATATCCTA	ATTCTAAGCA	GTACGTTGCG	3120
	GGTATTCACT	TGTCCTTC					3138

(2) INFORMATION FOR SEQ ID NO:8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2583 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...2583

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8

	TTAGAGATGG CATACGACTT TACACAAACA TTCCGCAACA GCCTGGAGTA CAGCTATCAG	60
25	GAAGCAACCC GTCTCGGCGT CGTAGCCGTG ACGCAAGATA TGCTCGTACT CGGTATCATT	120
	CGCGACGGAG ACAATGGCGC GATCGACATC ATGCGGCACT ATGGGATCAA CTTGTACGAA	180
	CTCAAACGGT TGATCGAGTT GGAAGCCATC GCCGAGAGTT TGCCTGCTTC GCCTGAGGGA	240
	TCGCCCATCT TCACCCCTTC GGCTCGGGAG GCTATCGATG ATGCCACAGA CATCTGTGCC	300
	GACATGGAGG ACGAGGCCGT CAGCCCGGTC CATCTGTTGC TGAGTATCCT CAACTCGACA	360
30	CAGGAGAGCT TAGTACAAAA GATATTTATG AAACAAGGTA TAAAATACGA CACCATCCTG	420
	TCGGATTACT TCGGACAGCG CAACCCCTCC GAAGGGAAGT CTCCCTCCGA AATGGAGATC	480
	CTCGACGGGT ACCAAGACAA CGACTTCGAC GACGAAGAGG ACGAATCCTC TCCGCCTTCC	540
	GGGAATAGCG GGACAGGCGG AGGCTCCGGC GACGCCCCCG AACAGAATAC CGGCGGAGGC	600
	GATACTACCA CCACGACACG GAGTGGAGGC GACACGCCTG CACTGGACAC CTTGCGCACC	660
35	GACATCACTG CCATGGCGGC AGCAGGCAAG CTCGACCCGG TAGTGGGTCG GGAGCAGGAG	720
	ATCGAAAGGG TGATACAGAT ACTCAGCCGG CGCAAAAAGA ACAATCCGGT GCTCATCGGC	780
	GAACCCGGTG TAGGCAAGAG TGCCATCGTG GAAGGACTGG CCGAACGCAT CGTGAACAGG	840
	AAGGTGAGCC GTATTCTTTT CGACAAGCGG ATCATCAGCC TCGATTGCGC TCAGATGGTA	900
	GCCGGCACCA AATATCGCGG ACAGTTCGAA GAGCGGTTGA AAGCCGTGCT CGATGAGCTG	960
40	AAGAAGAATC CGCAGATCAT CCTCTTCATC GACGAGATAC ATACCATCGT GGGAGCAGGC	1020

	TCTGCAGCCG	GATCGATGGA	TACGGCCAAT	ATGCTCAAAC	CCGCTCTTGC	CCGTGGACAG	1080
	GTACAGTGCA	TCGGAGCCAC	TACGCTGGAT	GAGTATCGTA	AGAACATAGA	AAAGGACGGA	1140
	GCACTCGAAC	GCCGCTTCCA	GAAGGTGCCG	ATAGCCCCCT	CGACTGCAGA	AGAAACGCTG	1200
	ACCATCCTGC	AAAACATCAA	AGAGAAATAC	GAGGACTATC	ACGGTGTACG	CTATACGGAC	1260
5	GAAGCGATCA	AAGCGGCAGT	GGAAGTGACC	GATCGCTATG	TATCCGATCG	TTTCTTCCCA	1320
	GATAAGGCGA	TAGATGCCAT	GGACGAGGCC	GGCGCGAGCG	TCCATATCAC	CAATGTGGTG	1380
	GCTCCGAAAG	AAATCGAGAT	ACTGGAGGCC	GAATTGGCAT	CGGTGCGAGA	GAACAAGCTC	1440
	TCGGCCGTAA	AGGCTCAGAA	CTACGAACTG	GCTGCCTCCT	TCCGCGATCA	GGAGCGGCGC	1500
	ACTCAGCAGC	AGATAGCGGA	AGAGAAGAAA	AAATGGGAAG	AGCAGATGTC	CAAGCACCGC	1560
10	GAGACGGTGG	ACGAGAATGT	AGTGGCGCAT	GTAGTGGCGT	TGATGACAGG	CGTTCCGGCT	1620
	GAGCGGCTGA	GCACGGGCGA	AGGCGAACGT	CTGCGCACGA	TGGCAGATGA	TCTCAAGACC	1680
	AAAGTAGTAG	GTCAGGACAC	AGCCATCGAA	AAGATGGTGC	ATGCCATCCA	GCGCAATCGT	1740
	CTGGGACTTC	GCAATGAAAA	GAAACCGATC	GGTTCTTTCC	TTTTCTTCGG	CCCCACGGGG	1800
	GTAGGCAAGA	CCTATTTGGC	CAAGAAGCTC	GCCGAATACC	TGTTTCGAGGA	TGAGAATGCC	1860
15	ATGATCAGGG	TGGATATGAG	CGAGTATATG	GAGAAGTTCT	CCGTTTCGCG	TCTCGTGGGT	1920
	GCCCCCTCCG	GATATGTGGG	CTATGAAGAA	GGCGGCCAAC	TGACGGAGCG	CGTAAGACGC	1980
	AAACCCTATT	CCGTGGTTCT	CTTGGATGAG	ATCGAAAAGG	CGCATGCCGA	TGTCTTCAAT	2040
	CTGCTCTTAC	AGGTGATGGA	CGAAGGTCAG	CTGACCGACA	GTCTGGGACG	GCGCGTGAAT	2100
	TTCAAGAACA	CCGTGATCAT	CATCACCTCC	AACGTGGGTA	CACGCCAGCT	CAAAGACTTC	2160
20	GGGCAGGGTA	TCGGGTTCG	TTCGGA AAAA	GACGAGGAAG	CGAACAAGGA	GCATAGCCGT	2220
	TCCGTGATCC	AAAAAGCTCT	GAACAAGACG	TTCAGCCCCG	AATTTCTCAA	CCGTTTGGAC	2280
	GATATCATCC	TCTTCGACCA	ACTGGGCAAG	ACGGAGATTG	GCCGGATGGT	GGACATAGAG	2340
	CTTAAAGCCG	TCTTGGCGCG	CATCCATCGT	GCCGGATACG	ACCTCGTCCT	TACCGATGAA	2400
	GCCAAGGATG	TGATAGCGAC	GAAGGGATAC	GACCTCCaAT	ACGGAGCACG	ACCGCTCAAG	2460
25	CGCACACTCC	AGAACGAAGT	GGAGGATCGC	CTCACGGATC	TTATCCTCTC	CGGACAGATC	2520
	GAGAAAGGGC	AGACGCTTAC	GCTCTCTGCT	CGCGATGGCG	AGATCATCGT	ACAAGAACAA	2580
	GCA						2583

30 (2) INFORMATION FOR SEQ ID NO:9

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 897 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...897

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9

	TTAAAAAAGG AGATAACTAT GAAACAGAAC TACTTCAAAA GAGTCTGCTC ACTGCTTTGG	60
	CTGGTTTTAC CCATGCTTAT TATGCCATTG GAAGTAGCAG CTCAAGAGAT TATTCCGAAC	120
15	GAAGAGGTGT TGGAATCATT GACTTTCGTT GCACCGGTTG AGGAGACAGA CGCAATAGAG	180
	GCAGAGGTAG AAGCTCTGCA GGAGATAGTC GCTACTGAGG AGATTGCGGA GCAGGCTGTT	240
	CGTTCCTTATA CCTACACGGT CTATCGTGAT GCGGTGAAGA TTGCTTCAGG ATTGACTGAG	300
	CCCACTTTTC TCGATGAAGA TGTTCTTGCC GGCGAACATA CCTACTGCGT AGAAGTACAG	360
	TATCAGGGAG GCGTATCCGA CAAAGTATGC GTGGACGTAG AGGTGAAGGA CTTCAAACCG	420
20	GTTACCAATC TCACCGGAAC TGCTTCCAAT GACGAAGTTT CTTTGGACTG GGACGGTGTG	480
	GAAGAGAAAG CTGAAGAGCC GGCAAGTGAT AAAGCAGTCA GCTACAACGT CTACAAGAAT	540
	GGAACCTTGA TCGGTAATAC AGCTGAAACT CATTATGTGG AGACCGGTGT AGCCAATGGT	600
	ACATACATCT ACGAAGTGGA AGTAAAGTAT CCTGACGGTG TATCTCCGAA GGTGGCTGTA	660
	ACCGTGACCG TGACCAACAG CTCATTGAGC AATGTAGATG GACAGGCTCC TTACACATTG	720
25	CGAGTAGAAG GCAAGAAGAT TATTGCGGAA GCCCATGGTA TGATCAGCT CTACGACATC	780
	AACGGACGTA CCGTGCCGT AGCCCCGAAT CGATTGGAAT ACATGGCGCA AACCGGTTTC	840
	TATGCAGTGC GCTTCGATGT GGGGAATAAA CACCATGTAT CGAAAATACA AGTAAGA	897

30 (2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1392 base pairs

(B) TYPE: nucleic acid

35 (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

40 (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1392

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10

	TACGGCAAAA	GAAGAAAGCT	CGGAACCTCC	GTCCGACCTT	CCGTTCTAAC	CCAAATCAGA	60
	TTTATTTTGG	ACTTACACCT	TATTACTGAT	TTTTTTGAAG	GGCTTCGGGT	CAATCCTATC	120
15	GGTGCAGCAG	CCATAGTGGC	TTTCATTATC	GACCTGCTTC	TTCTTTGCTG	TTCGGCTTTT	180
	ATGTCCTCCT	GTGAGGTGGC	TTATTTTTCA	CTAAAGCCGA	TCGATCTGCA	GAACATCCGC	240
	GAACGGAATC	ACTCTTCCGA	CATCGCGCTT	TCCAATTTAT	TAGACAATTC	GAATCAGCTA	300
	TTAGCTACTA	TTCTGATCGG	GAATAATGTG	ATTAATGTAG	CCATCGTTAT	CCTTTCCAAT	360
	TATGCCATCG	AGCAGACATT	CGTTTTCTCT	TCTCCGATCA	TTGGATTTCT	GATCCAGACG	420
20	ATACTCCTGA	CCACTGTTCT	TTTGCTGTTC	GGAGAGATTC	TGCCGAAAGT	GTATGCGCGG	480
	AAGAATCCGC	TGCAATACTC	GCGCTTTTCT	GCTGCAGCTA	TGTCCGTTAT	CTATAAGATA	540
	TTGTCACCGT	TTTCAAATTT	GCTGGTCAAA	AGTACCGGCA	TCGTTACCAG	AGGTATCAGC	600
	AAGAAGAAAT	ACGATATGTC	CGTGGATGAG	CTCTCGAAAG	CGGTAGCCCT	CACCACTACG	660
	GAGGGAGAGC	CGGAGGAGAA	AGAAATGATT	AACGAAATCA	TCAAATTCTA	TAATAAGACA	720
25	GCCTGCGAAA	TCATGGTTCC	GCGTATCGAT	ATTGTGGATG	TGGATCTGAG	CTGGCCATTT	780
	CGTAAGATGC	TTGACTTCGT	TGTTTCGTCG	GGTTATTCCA	GACTTCCCGT	TTCAGAGGGG	840
	TCAGAAGACA	ATATCAAAGG	GGTGATTTAC	ATCAAAGATC	TAATCCCACA	CATGGATAAA	900
	GGCGATGAAT	TCGACTGGCA	TCCTCTGATT	CGTAAAGCAT	ATTTTGTCCC	CGAAAACAAG	960
	CGCATAGATG	ATTTGCTCGA	GGAGTTCAGA	GCCAATAAGG	TGCATGTCTC	CATCGTTGTG	1020
30	GATGAGTTCG	GTGGCACTTG	CGGACTGATC	ACAATGGAGG	ACATATTGGA	AGAGATCGTC	1080
	GGCGAGATTA	CGGACGAGTA	CGATGAGGAA	GAACTCCCCT	TTAAGGTTTT	GGGGGATGGC	1140
	AGTTATCTTT	TCGAAGGAAA	AACGTCTCTC	TCCGATGTTT	GACACTATCT	TGACCTTCCG	1200
	GAAAATGCTT	TCGGTGAATT	GGGGGACGAG	GTAGATACGC	TAAGTGGGCT	CTTCTTGGA	1260
	ATCAAGCAGG	AACTCCCCCA	TGTGGGCGAT	ACAGCAGTGT	ACGAGCCATT	CCGCTTTCAA	1320
35	GTGACCCAAA	TGGACAAGCG	CCGAATCATC	GAAATCAAGA	TTTTCCCTTT	CGAGCGCACT	1380
	TGGGAGGTCG	AA					1392

(2) INFORMATION FOR SEQ ID NO:11

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2607 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...2607

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11

TGGCATAGGA ATATTTTTAT CTTTGCGAGT ACATTTAGCC CGAAAAATAT GCTCCCACTG	60
CCATACCGTT ATGCAAAAAC CGAGCACCTT TTTCTCGCAA AAGGATACTG CAAGAATCCA	120
ATAACAAACA TAATTATCCT ATTTATGAAG AAAAAGAATT TTTTGCTTCT TGGCATTTC	180
GTTTGCTTTGC TGACTTTCAT CGGCAGCATG CAGGCACAAC AGGCCAAAAGA TTATTTCAAC	240
TTTGACGAAC GGGGCGAGGC CTACTTCTCA TTCAAAGTGC CTGATAGGGC CGTTCACAA	300
GAGCTGGCTC TGATCATGTC CATCGACGAG TTTGACCCCG TAACCAATGA AGCCATTGCC	360
TATGCCAGCG AAGAGGAGTT CGAGGCATTC CTGCGCTATG GGCTCAAGCC TACATTCTTG	420
ACTCCTCCAT CCATGCAGCG CGCTGTCGAG ATGTTGCGACT ACCGCTCAGG AGAAAAATAC	480
GAATGGAATG CTTACCCAC CTATGAAGCC TATATCAGCA TGATGGAAGA GTTCCAAACA	540
AAGTATCCAT CACTTTGTAC TACTTCCGTC ATTGGCAAGT CCGTAAAGGA TCGTAAACTG	600
ATGATTTGCA AGCTGACGTC CTCTGCCAAT ACAGGGAAAA AGCCTCGCGT GCTCTATACT	660
TCTACGATGC ACGGAGACGA AACGACCGGA TATGTGGTAC TGCTCCGACT CATAGACCAT	720
CTGCTGTCGA ACTACGAATC CGATCCGAGG ATTAAGAACA TTCTGGATAA AACGGAAGTA	780
TGGATCTGCC CTTTGACCAA TCCGGACGGA GCATACAGAG CCGGAAACCA CACCGTACAA	840
GGAGCTACTC GCTACAATGC CAACAATGTC GATTTGAACC GTAAC TTCAA GGATGATGTA	900
GCCGGTGATC ACCCCGATGG AAAACCTTGG CAGCCGGAGG CAACTGCATT CATGGATTTG	960
GAAGGAAACA CCTCTTTCTG GCTCGGTGCC AATATACATG GAGGAACAGA GGTGGTGAAC	1020
TATCCATGGG ATAATAAAAA AGAAAGACAT GCAGACGATG AGTGGTACAA ACTGATCAGT	1080
CGCAACTACG CAGCCGCTTG TCAGAGTATT TCCGCCAGCT ACATGACCTC CGAAACCAAT	1140

	TCGGGAATCA	TCAACGGTTC	AGACTGGTAT	GTAATTCGCG	GAAGTCGTCA	GGACAATGCA	1200
	AATTATTTCC	ATCGTCTGCG	AGAAATTACC	CTTGAAATCA	GCAACACGAA	GTTGGTGCCG	1260
	GCCTCTCAAC	TTCCAAAGTA	TTGGAATCTG	AACAAAGAAT	CTCTGCTTGC	TCTGATCGAA	1320
	GAATCCTTAT	ACGGCATCCA	TGGTACAGTG	ACTTCCGCTG	CGAACGGACA	GCCTCTCAAA	1380
5	TGCCAGATCT	TGATAGAAAA	CCATGACAAG	CGCAACTCCG	ATGTTTACTC	CGATGCTACC	1440
	ACAGGCTACT	ACGTACGTCC	TATCAAAGCC	GGCACTTATA	CGGTGAAATA	CAAAGCCGAG	1500
	GGTTATCCTG	AGGCAACTCG	TACCATTACG	ATCAAGGACA	AAGAAACCGT	CATCATGGAC	1560
	ATTGCATTGG	GCAACTCGGT	TCCTCTGCCT	GTACCCGATT	TCACAGCTTC	TCCTATGACC	1620
	ATCTCAGTAG	GCGAAAGCGT	CCAATTCCAA	GATCAAACGA	CAAATAACCC	CACGAATTGG	1680
10	GAGTGGACGT	TCGAAGGCGG	ACAGCCTGCC	ATGAGTACAG	AGCAGAATCC	GCTCGTATCC	1740
	TATAGTCATC	CCGGTCAGTA	CGACGTTACG	CTCAAAGTGT	GGAATGCAAG	TGGTTCCAAC	1800
	ACGATTACGA	AAGAAAAATT	CATCACTGTC	AATGCCGTTA	TGCCTGTAGC	TGAATTCGTC	1860
	GGTACCCCGA	CGGAAATAGA	AGAGGGCCAG	ACGGTATCTT	TCCAAAACCA	ATCCACCAAT	1920
	GCCACCAACT	ACGTATGGAT	ATTGATGGC	GGCACTCCCG	CTACCAAGTGA	AGACGAAAAC	1980
15	CCGACTGTGC	TTTACAGCAA	AGCCGGCCAA	TACGATGTCA	CGCTCAAGGC	GATCAGTGCT	2040
	TCCGGTGAAA	CGGTGAAGAC	GAAAGAAAAA	TACATCACTG	TCAAGAAAGC	TCCGGTCCCT	2100
	GCTCCGGTAG	CCGACTTCGA	AGGAACACCT	CGAAAAGTAA	AGAAAGGCGA	GACAGTTACT	2160
	TTCAAAGACT	TGTCTACGAA	CAATCCGACT	TCATGGCTTT	GGGTGTTTCA	AGGCGGCTCT	2220
	CCTGCCACCA	GCACGGAGCA	AAACCCGGTG	GTCACCTACA	ATGAAACAGG	CAAGTACGAT	2280
20	GTCCAGCTGA	CTGCCACCAA	CGAGGGCGGA	AGCAATGTGA	AGAAAGCAGA	AGACTACATT	2340
	GAGGTTATCC	TCGATGACAG	TGTCGAGGAC	ATAGTGGCAC	AGACGGGTAT	CGTCATTTCGT	2400
	CCGCAAAACG	GAACGAAGCA	GATCCTCATA	GAAGCCAACG	CTGCTATCAA	AGCGATCGTT	2460
	CTCTATGACA	TCAATGGACG	GGTCGTACTC	AAAATACTC	CGAATCAGCT	CCGCTCGACC	2520
	GTAGATCTTT	CCATCCTGCC	CGAAGGAATC	TACACCATCA	ATATCAAAAC	GGAAAAATCC	2580
25	GCTCGCACGG	AAAAGATCCA	TATCGGG				2607

(2) INFORMATION FOR SEQ ID NO:12

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 849 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

40 (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...849

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

	Phe	Gly	Ile	Ser	Pro	Ser	Met	Lys	Lys	Ser	Phe	Leu	Leu	Ala	Ile	Val
	1				5					10					15	
10	Met	Leu	Phe	Gly	Ile	Ala	Met	Gln	Gly	His	Ser	Ala	Pro	Val	Thr	Lys
				20					25					30		
	Glu	Arg	Ala	Leu	Ser	Leu	Ala	Arg	Leu	Ala	Leu	Arg	Gln	Val	Ser	Leu
				35				40					45			
	Arg	Met	Gly	Gln	Thr	Ala	Val	Ser	Asp	Lys	Ile	Ser	Ile	Asp	Tyr	Val
15		50					55					60				
	Tyr	Arg	Gln	Gly	Asp	Ala	Glu	Arg	Gly	Ile	Thr	Ser	Gln	Glu	Glu	Gly
	65				70					75					80	
	Ser	Pro	Ala	Tyr	Phe	Tyr	Val	Ala	Asn	Arg	Gly	Asn	Asn	Glu	Gly	Tyr
				85						90					95	
20	Ala	Leu	Val	Ala	Ala	Asp	Asp	Arg	Ile	Pro	Thr	Ile	Leu	Ala	Tyr	Ser
				100					105					110		
	Pro	Ile	Gly	Arg	Phe	Asp	Met	Asp	Ser	Met	Pro	Asp	Asn	Leu	Arg	Met
				115				120					125			
	Trp	Leu	Gln	Ile	Tyr	Asp	Gln	Glu	Ile	Gly	Leu	Ile	Leu	Ser	Gly	Lys
25		130					135					140				
	Ala	Gln	Leu	Asn	Glu	Glu	Ile	Leu	Arg	Thr	Glu	Gly	Val	Pro	Ala	Glu
	145				150					155					160	
	Val	His	Ala	Leu	Met	Asp	Asn	Gly	His	Phe	Ala	Asn	Asp	Pro	Met	Arg
				165					170					175		
30	Trp	Asn	Gln	Gly	Tyr	Pro	Trp	Asn	Asn	Lys	Glu	Pro	Leu	Leu	Pro	Asn
				180				185					190			
	Gly	Asn	His	Ala	Tyr	Thr	Gly	Cys	Val	Ala	Thr	Ala	Ala	Ala	Gln	Ile
				195				200					205			
	Met	Arg	Tyr	His	Ser	Trp	Pro	Leu	Gln	Gly	Glu	Gly	Ser	Phe	Asp	Tyr
35		210					215					220				
	His	Ala	Gly	Ser	Leu	Val	Gly	Asn	Trp	Ser	Gly	Thr	Phe	Gly	Glu	Met
	225				230					235					240	
	Tyr	Asp	Trp	Ile	Asn	Met	Pro	Gly	Asn	Pro	Asp	Leu	Asp	Asn	Leu	Thr
				245						250					255	
40	Gln	Ser	Gln	Val	Asp	Ala	Tyr	Ala	Thr	Leu	Met	Arg	Asp	Val	Ser	Ala

		260		265		270										
	Ser	Val	Ser	Met	Ser	Phe	Tyr	Glu	Asn	Gly	Ser	Gly	Thr	Tyr	Ser	Val
		275						280					285			
	Tyr	Val	Val	Gly	Ala	Leu	Arg	Asn	Asn	Phe	Arg	Tyr	Lys	Arg	Ser	Leu
5		290					295					300				
	Gln	Leu	His	Val	Arg	Ala	Leu	Tyr	Thr	Ser	Gln	Glu	Trp	His	Asp	Met
	305					310					315				320	
	Ile	Arg	Gly	Glu	Leu	Ala	Ser	Gly	Arg	Pro	Val	Tyr	Tyr	Ala	Gly	Asn
				325						330					335	
10	Asn	Gln	Ser	Ile	Gly	His	Ala	Phe	Val	Cys	Asp	Gly	Tyr	Ala	Ser	Asp
				340						345				350		
	Gly	Thr	Phe	His	Phe	Asn	Trp	Gly	Trp	Gly	Gly	Val	Ser	Asn	Gly	Phe
				355				360						365		
	Tyr	Lys	Leu	Thr	Leu	Leu	Ser	Pro	Thr	Ser	Leu	Gly	Ile	Gly	Gly	Glu
15		370					375					380				
	Gly	Ile	Gly	Phe	Thr	Ile	Tyr	Gln	Glu	Ile	Ile	Thr	Gly	Ile	Glu	Pro
	385					390					395				400	
	Ala	Lys	Thr	Pro	Ala	Glu	Ala	Gly	Thr	Asp	Ala	Leu	Pro	Ile	Leu	Ala
				405						410				415		
20	Leu	Lys	Asp	Ile	Glu	Ala	Glu	Tyr	Lys	Ser	Glu	Ser	Gly	Leu	Asn	Val
				420					425					430		
	Gly	Tyr	Ser	Ile	Tyr	Asn	Thr	Gly	Glu	Glu	Gln	Ser	Asn	Leu	Asp	Leu
				435				440					445			
	Gly	Tyr	Arg	Leu	Asn	Lys	Ala	Asp	Gly	Glu	Val	Ile	Glu	Val	Lys	Thr
25		450					455					460				
	Ser	Ser	Ile	Asn	Ile	Ser	Trp	Tyr	Gly	Tyr	Gly	Glu	His	Pro	Glu	Ser
	465					470				475				480		
	Phe	Ser	Leu	Ala	Pro	Asn	Gln	Leu	Ser	Gln	Gly	Ile	Asn	Thr	Ile	Thr
				485					490					495		
30	Leu	Leu	Tyr	Arg	Arg	Thr	Gly	Thr	Glu	Gln	Trp	Glu	Pro	Val	Arg	His
				500					505					510		
	Ala	Gln	Gly	Gly	Tyr	Val	Asn	Ser	Ile	Lys	Val	Asn	Thr	Thr	Asp	Pro
				515					520				525			
	Asn	Asn	Val	Val	Val	Thr	Val	Asp	Asn	Asn	Glu	Gly	Lys	Leu	Ser	Ile
35		530					535					540				
	Val	Pro	Asn	Ser	Phe	Val	Ala	Asp	Leu	Asn	Ser	Tyr	Glu	His	Ser	Thr
	545					550				555				560		
	Ile	Thr	Val	Gln	Phe	Asn	Ser	Asp	Ser	Pro	Asp	Glu	Ile	Arg	Thr	Pro
				565					570					575		
40	Val	Ala	Phe	Ala	Leu	Ser	Thr	Gly	Ala	Thr	Ala	Asp	Asp	Val	Ile	Ser

	580		585		590
	Leu Gly Trp Val Met Ala Glu Val Pro Gly Gly Ser Ser Asn Tyr Pro				
	595		600		605
	Val Val Trp Ser Lys Asp Val Leu Thr Leu Ser Glu Gly Asp Tyr Thr				
5	610		615		620
	Leu Trp Tyr Arg Phe Ser Ile Asn Asn Gln Lys Asp Glu Trp Lys Lys				
	625		630		635
	Ile Gly Ser Val Ser Val Lys Thr Pro Thr Glu Tyr Thr His Pro Leu				
		645		650	655
10	Phe Glu Val Gly His Asn Gln Thr Ser Thr Tyr Thr Leu Asp Met Ala				
		660		665	670
	His Asn Arg Val Leu Pro Asp Phe Thr Leu Lys Asn Leu Gly Leu Pro				
		675		680	685
	Phe Asn Gly Glu Leu Val Val Val Phe Arg Gln Thr Gln Ser Ser Ser				
15	690		695		700
	Gly Ser Leu Trp Ala Ala Gln Glu Thr Val His Ile Lys Gln Gly Glu				
	705		710		715
	Thr Phe Val Tyr Lys Pro Val Val Glu Gly Pro Ile Pro Asp Gly Ser				
		725		730	735
20	Tyr Arg Ala Thr Leu His Ala Phe Val Asn Gly Gln Gln Gln Leu Tyr				
		740		745	750
	Leu Lys Gly Lys Arg Asn Tyr Thr Val Lys Ile Val Asn Gly Thr Ala				
		755		760	765
	Val Glu Ala Ile Glu Ser Ser Glu Glu Ile Arg Val Phe Pro Asn Pro				
25	770		775		780
	Ala Arg Asp Tyr Val Glu Ile Ser Ala Pro Cys Ile Pro Gln Glu Thr				
	785		790		795
	Ser Ile Ile Leu Phe Asp Leu Ser Gly Lys Ile Val Met Lys Asn Ser				
		805		810	815
30	Leu Ser Ala Gly His Gly Arg Met Asp Val Ser Arg Leu Pro Asn Gly				
		820		825	830
	Ala Tyr Ile Leu Lys Val Asp Gly Tyr Thr Thr Lys Ile Asn Ile Val				
		835		840	845
	His				
35					

(2) INFORMATION FOR SEQ ID NO:13

(i) SEQUENCE CHARACTERISTICS:

40 (A) LENGTH: 1269 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Porphyromonas gingivalis

10

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1269

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

	Val	Lys	Cys	Met	Gly	Lys	Tyr	Lys	Arg	Ala	Lys	Tyr	Arg	Tyr	Trp	Leu	
	1				5					10					15		
	Phe	Pro	Phe	Cys	Ser	Asp	Tyr	Tyr	Thr	Phe	Glu	Gly	Val	Thr	Phe	Leu	
20				20					25					30			
	Cys	Ala	Ser	Asp	Asp	Met	Thr	Thr	Lys	Lys	Pro	Gln	Ala	Ile	Leu	Asp	
				35					40					45			
	Leu	Glu	Lys	Ala	Tyr	Asn	Ile	Glu	Ile	Pro	Asp	Leu	Ser	Ser	Gln	Glu	
				50				55						60			
25	Gly	Ile	Ser	Trp	Ser	Val	Asn	Arg	Tyr	Phe	Lys	Gln	Asp	Ser	Ser	Gly	
	65					70					75					80	
	Ala	Val	Val	Glu	Leu	Cys	Leu	Arg	Glu	Cys	Gln	Ile	Glu	Ser	Met	Thr	
						85				90					95		
	Trp	Leu	Ile	Asp	Phe	Pro	Ala	Leu	Lys	Lys	Leu	Asp	Leu	Ser	Tyr	Asn	
30						100				105					110		
	Gln	Ile	Ser	Lys	Leu	Glu	Gly	Leu	Glu	Arg	Leu	Thr	Ser	Leu	Thr	Lys	
						115				120					125		
	Leu	Arg	Leu	Arg	Ser	Asn	Gln	Ile	Arg	Lys	Leu	Glu	Gly	Leu	Asp	Ser	
						130				135					140		
35	Leu	Thr	Ser	Leu	Thr	Lys	Leu	Ser	Leu	Ser	Asp	Asn	Gln	Ile	Ser	Lys	
	145					150					155					160	
	Leu	Glu	Gly	Leu	Glu	Arg	Leu	Thr	Ser	Leu	Ala	Glu	Leu	Tyr	Leu	Leu	
						165					170				175		
	Asp	Asn	Gln	Ile	Ser	Lys	Leu	Glu	Gly	Leu	Glu	Arg	Leu	Thr	Ser	Leu	
40						180					185					190	

	Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu Gly Leu	
	195	200 205
	Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile	
	210	215 220
5	Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys Leu Arg	
	225	230 235 240
	Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr	
		245 250 255
	Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu	
10		260 265 270
	Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn	
		275 280 285
	Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Ser Ser Leu Thr Lys	
		290 295 300
15	Leu Arg Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg	
	305	310 315 320
	Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile Ser Lys	
		325 330 335
	Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr Leu Leu	
20		340 345 350
	Asp Asn Gln Ile Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu	
		355 360 365
	Thr Lys Leu Arg Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu	
		370 375 380
25	Asp Ser Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile	
	385	390 395 400
	Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr	
		405 410 415
	Leu Leu Asp Asn Gln Ile Arg Lys Leu Glu Gly Leu Asp Gly Leu Ala	
30		420 425 430
	Ser Leu Thr Arg Leu Ser Leu Arg Arg Asn Gln Ile Ser Lys Leu Glu	
		435 440 445
	Gly Leu Asp Arg Leu Lys Val Leu Arg Lys Leu Asp Val Ser Gly Asn	
		450 455 460
35	Asp Ile Gln Ser Ile Asp Asp Ile Lys Leu Leu Ala Pro Ile Leu Glu	
	465	470 475 480
	Gln Thr Leu Glu Lys Leu Arg Ile His Asp Asn Pro Phe Val Ala Ser	
		485 490 495
	Ser Gly Leu Ile Leu Ser Pro Tyr Asp Asn His Leu Pro Glu Ile Lys	
40		500 505 510

40

	His	Asn	Leu	Ser	Ser	Gly	Ser	Ile	Phe	Glu	Glu	Asp	Gly	Gln	Asn	Gly	
			835					840					845				
	Asn	Met	Ile	Leu	Gln	Leu	Leu	Leu	Glu	Glu	Leu	Ile	Val	Tyr	Glu	Asp	
			850				855					860					
5	Lys	Asp	Cys	Tyr	Val	Ile	Pro	Gly	Tyr	Leu	Pro	Leu	His	Ser	Asp	Asp	
			865			870				875						880	
	Glu	Ala	Tyr	Lys	Trp	Leu	Thr	Leu	Gly	Phe	Glu	Arg	Pro	Asn	Phe	Val	
					885				890						895		
	Leu	Lys	Phe	Glu	Arg	Phe	Ile	Pro	Phe	Gly	Leu	Ile	Asn	Gln	Ile	Ile	
10				900					905					910			
	Ala	Tyr	Tyr	Gly	Arg	Glu	Glu	Gly	Ala	Leu	Lys	Arg	Tyr	Trp	Arg	Asp	
				915				920					925				
	Gln	Val	Ile	Phe	Thr	Ala	Gly	Arg	Glu	Met	Asp	Arg	Gln	Thr	Leu	Glu	
				930			935				940						
15	Gln	Glu	Glu	Glu	Lys	Glu	Gly	Leu	Pro	Lys	Thr	Asn	Ala	Glu	Asp	Tyr	
						945		950			955					960	
	Gln	Ile	Trp	Ile	Lys	Leu	Asp	Phe	Thr	Asp	Leu	Ala	Ile	Ser	Val	Phe	
					965				970						975		
	Ile	Lys	Glu	Gln	Arg	Lys	Thr	Ser	Ala	Lys	Asp	Met	Gln	Arg	Lys	Glu	
20				980					985					990			
	Ala	Thr	Ile	Leu	Ser	Asp	Met	Leu	Asp	Met	Tyr	Trp	Asn	Asn	Ile	Pro	
				995				1000					1005				
	Pro	Arg	Glu	Gln	Ile	Gly	Asp	Lys	Asp	Thr	Glu	Gln	Thr	Arg	Ser	Thr	
				1010			1015				1020						
25	Ile	Arg	Glu	Thr	Asn	Arg	Lys	Lys	Arg	Pro	Ile	Gln	Asp	Leu	Tyr	Leu	
				1025			1030				1035					1040	
	Ser	Cys	Ala	Gln	Ala	Asp	Lys	Asp	Leu	Thr	Glu	Ser	His	Tyr	Ile	His	
					1045				1050					1055			
	Leu	Gly	Thr	Leu	Asp	Asp	Glu	Ser	Lys	Thr	Thr	Ala	Arg	Ile	Ala	Ala	
30				1060					1065					1070			
	Tyr	Pro	Leu	Lys	Asn	Gly	Val	Ile	Asp	Lys	Glu	Arg	Val	Arg	Glu	Val	
				1075			1080				1085						
	Ser	Thr	Arg	Pro	Tyr	Lys	His	Leu	Ser	Val	Asn	Lys	Asn	Leu	Ala	Thr	
				1090			1095				1100						
35	Ala	Lys	Gln	Ile	Phe	Ile	Ser	Tyr	Ser	Lys	Glu	Asp	Gln	Thr	Glu	Leu	
				1105			1110				1115					1120	
	Glu	Thr	Cys	Leu	Gln	Phe	Phe	Lys	Pro	Leu	Glu	Lys	Asn	Gly	Gln	Ile	
					1125				1130					1135			
	Glu	Ile	Tyr	Tyr	Asp	Lys	Leu	Thr	Lys	Phe	Glu	Thr	Pro	Ile	His	Pro	
40				1140					1145					1150			

Glu Ile Arg Lys Arg Ile Val Glu Ala Asp Cys Ile Ile Ala Leu Ile
 1155 1160 1165
 Ser Gln Arg Tyr Leu Ala Thr Asp Tyr Ile Leu Asp His Glu Leu Pro
 1170 1175 1180
 5 Val Phe Arg Glu Tyr Asn Lys Thr Ile Val Pro Ile Leu Ile Lys Pro
 1185 1190 1195 1200
 Cys Thr Phe Glu Asp Asp Glu Phe Leu Arg Glu Lys Tyr Phe Ala Gln
 1205 1210 1215
 Lys Ala Gln Ile Ile Asn Leu Gly Lys Glu Gly Lys Thr Ile Lys Ala
 10 1220 1225 1230
 Tyr Asp Ser Ile Thr Ala Ser Ala His Arg Asp Glu Asn Trp Val Ala
 1235 1240 1245
 Val Val Arg Glu Phe Lys Glu Lys Ile Leu Arg Ile Thr Lys Gln Glu
 1250 1255 1260
 15 Val Asn Thr Asp Glu
 1265

(2) INFORMATION FOR SEQ ID NO:14

20 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

30 (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...377

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

Leu Phe Asn His Lys Lys Ser Trp Tyr Glu Thr Phe Gln Phe Tyr Leu
 1 5 10 15
 40 Val Val Phe Arg Ser Gly Phe Ile Leu Leu Cys Glu Asn Thr Leu Ala

	20				25				30							
	Gln	Gln	Lys	Thr	Glu	Glu	Phe	Ala	Pro	Val	Ser	Asp	Leu	Arg	Ala	Glu
	35				40				45							
	Ala	Tyr	Gly	Ser	Thr	Val	Phe	Leu	His	Trp	Thr	Pro	Pro	Tyr	Asp	Asn
5	50				55				60							
	Pro	Met	Ile	Pro	Leu	Ser	Glu	Ser	Phe	Glu	Ser	Gly	Ile	Pro	Ala	Ile
	65				70				75				80			
	Trp	Lys	Thr	Ile	Asp	Ala	Asp	Gly	Asp	Gly	Tyr	Asn	Trp	Met	His	Leu
	85				90				95							
10	Thr	Asn	Phe	Thr	Gly	Gln	Ser	Gly	Leu	Cys	Val	Ser	Ser	Ala	Ser	Tyr
	100				105				110							
	Ile	Gly	Gly	Val	Gly	Ala	Leu	Thr	Pro	Asp	Asn	Tyr	Leu	Ile	Thr	Pro
	115				120				125							
	Glu	Leu	Lys	Leu	Pro	Thr	Asp	Ala	Leu	Val	Glu	Ile	Ile	Tyr	Trp	Val
15	130				135				140							
	Cys	Thr	Gln	Asp	Leu	Thr	Ala	Pro	Ser	Glu	His	Tyr	Ala	Val	Tyr	Ser
	145				150				155				160			
	Ser	Ser	Thr	Gly	Asn	Asn	Ala	Ala	Asp	Phe	Val	Asn	Leu	Leu	Tyr	Glu
	165				170				175							
20	Glu	Thr	Leu	Thr	Ala	Lys	Arg	Ile	Gln	Ser	Pro	Glu	Leu	Ile	Arg	Gly
	180				185				190							
	Asn	Arg	Thr	Gln	Gly	Val	Trp	Tyr	Gln	Arg	Lys	Val	Val	Leu	Pro	Asn
	195				200				205							
	Asp	Thr	Lys	Tyr	Val	Ala	Phe	Arg	His	Phe	Asn	Ser	Thr	Asp	Asn	Phe
25	210				215				220							
	Trp	Leu	Asn	Leu	Asp	Glu	Val	Ser	Ile	Leu	Tyr	Thr	Pro	Leu	Pro	Arg
	225				230				235				240			
	Arg	Ala	Pro	Cys	Pro	His	Pro	Gly	Gly	Tyr	Thr	Tyr	Ser	Val	Phe	Arg
	245				250				255							
30	Asp	Gly	Gln	Lys	Ile	Ala	Ser	Gly	Leu	Ser	Ala	Leu	Ala	Tyr	Ile	Asp
	260				265				270							
	Thr	Asp	Val	Pro	Tyr	Gly	Thr	Gln	Asp	Tyr	Cys	Val	Gln	Val	Asn	Tyr
	275				280				285							
	Leu	Gln	Gly	Asp	Ser	Tyr	Lys	Val	Cys	Lys	Asn	Ile	Val	Val	Ala	Asn
35	290				295				300							
	Ser	Ala	Asn	Ile	Tyr	Gly	Ala	Asp	Lys	Pro	Phe	Ala	Leu	Thr	Val	Val
	305				310				315				320			
	Gly	Lys	Thr	Ile	Val	Ala	Ser	Ala	Phe	Lys	Gly	Glu	Ile	Thr	Leu	Tyr
	325				330				335							
40	Asp	Ile	Arg	Gly	Arg	Leu	Ile	Ala	Ser	Gly	Cys	Asp	Thr	Leu	Arg	Tyr

5

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

10

(ii) MOLECULE TYPE: protein

15

(vi) ORIGINAL SOURCE:

20

(ix) FEATURE:

(A) NAME/KEY: misc feature

25 : (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

30

35

40

	Arg	Thr	Asn	Asn	Leu	Glu	Glu	Val	Val	Val	Thr	Gly	Thr	Gly	Thr	Arg	
				115					120					125			
	Tyr	Arg	Leu	Val	Asp	Ala	Pro	Val	Ala	Thr	Glu	Val	Leu	Thr	Ala	Lys	
			130					135					140				
5	Asp	Ile	Ala	Ser	Phe	Ser	Ala	Pro	Thr	Ser	Glu	Ala	Leu	Leu	Gln	Gly	
	145						150					155				160	
	Leu	Ser	Pro	Ser	Phe	Asp	Phe	Gly	Pro	Asn	Leu	Met	Gly	Ser	Phe	Met	
					165					170					175		
	Gln	Leu	Asn	Gly	Leu	Ser	Ser	Lys	Tyr	Ile	Leu	Ile	Leu	Ile	Asp	Gly	
10				180						185					190		
	Lys	Arg	Val	Tyr	Gly	Asp	Val	Gly	Gly	Gln	Ala	Asp	Leu	Ser	Arg	Ile	
			195					200					205				
	Ser	Pro	Asp	Gln	Ile	Glu	Arg	Ile	Glu	Leu	Val	Lys	Gly	Ala	Ser	Ser	
			210				215						220				
15	Ser	Leu	Tyr	Gly	Ser	Asp	Ala	Ile	Ala	Gly	Val	Ile	Asn	Val	Ile	Thr	
	225					230					235					240	
	Lys	Lys	Asn	Thr	Asn	Arg	Leu	Ser	Ala	Tyr	Thr	Ser	His	Arg	Ile	Ser	
				245						250					255		
	Lys	Tyr	Asn	Asp	Arg	Gln	Thr	Asn	Thr	Ser	Leu	Asp	Ile	Asn	Ile	Gly	
20				260					265					270			
	Lys	Phe	Ser	Ser	Asn	Thr	Asn	Tyr	Phe	Phe	Tyr	His	Thr	Asp	Gly	Trp	
			275					280					285				
	Gln	Asn	Ser	Pro	Phe	Glu	Ile	Lys	Lys	Lys	Lys	Gly	Ser	Gly	Glu	Pro	
			290				295					300					
25	Val	Leu	Glu	Glu	Thr	Tyr	Lys	Lys	Thr	Phe	Arg	Ala	Gln	Glu	Asn	Gln	
	305						310				315					320	
	Gly	Val	Ser	Gln	Ser	Leu	Ser	Tyr	Tyr	Ala	Thr	Asn	Asn	Leu	Ser	Phe	
				325						330					335		
	Ser	Gly	Asn	Val	Gln	Tyr	Asn	Lys	Arg	Gln	Ile	Phe	Thr	Pro	Thr	Phe	
30				340					345					350			
	Ser	Glu	Lys	Lys	Ala	Tyr	Asp	Met	Asp	Tyr	Arg	Ala	Leu	Thr	Ala	Ser	
			355					360					365				
	Leu	Gly	Thr	Asn	Tyr	Leu	Phe	Pro	Asn	Gly	Leu	His	Thr	Leu	Ser	Phe	
			370				375					380					
35	Asp	Ala	Val	Tyr	Asp	Arg	Phe	Arg	Phe	Gly	Tyr	Leu	Tyr	His	Asp	Lys	
	385					390					395					400	
	Asp	Ser	Ser	Glu	Ser	Leu	Ile	Asn	Asn	Gln	Gly	Gln	Thr	Glu	Gln	Pro	
				405						410					415		
	Thr	Phe	Phe	Pro	Gly	Gln	Leu	Arg	Asn	Lys	Asn	Asp	Gln	Ile	Arg	Tyr	
40				420						425					430		

Thr Ala Glu Ala Arg Gly Val Phe Thr Leu Pro Tyr Ala Gln Lys Leu
 435 440 445
 Thr Gly Gly Leu Glu Tyr Phe Arg Glu Glu Leu Ile Ser Pro Tyr Asn
 450 455 460
 5 Leu Ile Thr Asp Lys Ala Asp Ala Ser Thr Leu Ser Ala Tyr Val Gln
 465 470 475 480
 Asp Glu Trp Lys Pro Leu Asp Trp Phe Asn Met Thr Ala Gly Phe Arg
 485 490 495
 Leu Val His His Gln Glu Phe Gly Thr Arg Met Thr Pro Lys Val Ser
 10 500 505 510
 Ile Leu Ala Lys Tyr Gly Pro Leu Asn Phe Arg Ala Thr Tyr Ala Asn
 515 520 525
 Gly Tyr Lys Thr Pro Thr Leu Lys Glu Leu Phe Ala Arg Asn Glu Leu
 530 535 540
 15 Thr Thr Met Gly Ser His Asn Leu Tyr Leu Gly Asn Ala Asp Leu Lys
 545 550 555 560
 Pro Gln Met Ser Asp Tyr Tyr Ala Leu Gly Leu Glu Tyr Asn Gln Gly
 565 570 575
 Pro Ile Ser Phe Ser Ala Thr Val Tyr Asp Asn Glu Leu Arg Asn Leu
 20 580 585 590
 Ile Ser Phe Met Asp Ile Pro Thr Ser Pro Glu His Glu Ala Gln Gly
 595 600 605
 Ile Lys Lys Thr Lys Gln Tyr Ala Asn Ile Gly Lys Ala Arg Ser Arg
 610 615 620
 25 Gly Leu Asp Val Leu Cys Asp Ala Ser Ile Gly Trp Gly Ile Lys Leu
 625 630 635 640
 Gly Ala Gly Tyr Ser Leu Val Glu Ala Lys Asn Leu Gln Thr Asp Glu
 645 650 655
 Trp Leu Glu Gly Ala Ala Arg His Arg Ala Asn Val His Ala Asp Trp
 30 660 665 670
 Val His Tyr Trp Gly Gln Tyr Arg Leu Gly Val Ser Leu Phe Gly Arg
 675 680 685
 Ile Gln Ser Glu Arg Tyr Tyr Lys Asp Gly Asn Ala Pro Asp Tyr Thr
 690 695 700
 35 Leu Trp Arg Leu Ala Thr Ser His Arg Phe Ala His Phe Arg His Ile
 705 710 715 720
 Ile Leu Asp Gly Thr Leu Gly Ile Asp Asn Leu Phe Asp Tyr Val Asp
 725 730 735
 Asp Arg Pro Met Gly Val Asn Tyr Ala Thr Val Thr Pro Gly Arg Thr
 40 740 745 750

Phe Phe Ala Gln Ile Ala Ile Arg Phe Asn Asn
 755 760

(2) INFORMATION FOR SEQ ID NO:16

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 582 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

15

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Porphyromonas gingivalis

(ix) FEATURE:

(A) NAME/KEY: misc_feature

20

(B) LOCATION 1...582

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

25	Ser	Lys	Ile	Val	Leu	Arg	Lys	Phe	Cys	Thr	Leu	Ala	Arg	Met	Lys	Lys
	1			5						10					15	
	Thr	Asn	Leu	Phe	Leu	Ser	Leu	Leu	Val	Ile	Phe	Ile	Thr	Gly	Ser	Phe
			20						25					30		
	Met	Thr	Ala	Cys	Ala	Gln	Lys	Ser	Lys	Thr	Asn	Lys	Leu	Thr	Glu	Glu
			35				40						45			
30	Asp	Arg	Ser	Arg	Asn	Glu	Tyr	Val	Gln	Ser	Met	Asp	Val	Leu	Ser	Asn
		50				55					60					
	Ile	Ile	Gly	Asn	Val	Arg	Leu	Tyr	Phe	Val	Asp	Thr	Ile	Ser	Ile	Lys
	65				70					75					80	
	His	Met	Thr	Arg	Arg	Gly	Ile	Asp	Ala	Met	Leu	Gly	Gly	Leu	Asp	Pro
35				85					90					95		
	Tyr	Thr	Glu	Tyr	Ile	Pro	Tyr	Glu	Glu	Met	Asp	Glu	Leu	Lys	Leu	Met
			100					105					110			
	Thr	Thr	Gly	Glu	Tyr	Ala	Gly	Val	Gly	Ala	Ile	Ile	Ser	Gln	Arg	Pro
		115					120					125				
40	Asp	Ser	Ala	Val	Ile	Ile	Gln	Arg	Pro	Met	Glu	Gly	Met	Pro	Ala	Asp

	130		135		140											
	Glu	Ala	Gly	Leu	Ile	Ala	Gly	Asp	Arg	Ile	Leu	Thr	Ile	Asp	Gly	Lys
	145			150			155				160					
	Asp	Phe	Arg	Lys	Ser	Thr	Thr	Pro	Lys	Val	Ser	Gln	Ala	Leu	Lys	Gly
5				165				170						175		
	Ile	Ala	Gly	Thr	Val	Ala	Lys	Val	Thr	Val	Met	Arg	Tyr	Gly	Glu	Thr
				180				185						190		
	Lys	Pro	Arg	Thr	Phe	Ser	Val	Lys	Arg	Gln	Lys	Val	Ile	Met	Asn	Ser
				195				200						205		
10	Val	Thr	Tyr	Ser	Gly	Met	Leu	Asp	Gly	Ser	Ile	Gly	Tyr	Ile	Arg	Leu
	210					215						220				
	Asn	Asn	Phe	Thr	Asp	Lys	Ser	Ala	Glu	Glu	Val	Arg	Thr	Ala	Leu	Leu
	225					230						235				240
	Asp	Leu	Arg	Asp	Lys	Gln	Gly	Ala	Lys	Gly	Leu	Ile	Leu	Asp	Leu	Arg
15				245						250				255		
	Gly	Asn	Gly	Gly	Gly	Leu	Met	Gln	Ala	Ala	Ile	Glu	Ile	Val	Asn	Leu
				260						265				270		
	Phe	Val	Pro	Lys	Gly	Lys	Glu	Val	Val	Thr	Thr	Lys	Gly	Arg	Ile	Ala
				275						280				285		
20	Glu	Ser	Ala	Ser	Val	Phe	Arg	Thr	Leu	Thr	Glu	Pro	Ile	Asp	Thr	Lys
	290						295						300			
	Leu	Pro	Ile	Val	Val	Leu	Ile	Asp	Gly	Gln	Ser	Ala	Ser	Ser	Ser	Glu
	305					310						315				320
	Ile	Val	Ala	Gly	Ala	Leu	Gln	Asp	Met	Asp	Arg	Ala	Val	Leu	Met	Gly
25				325						330				335		
	Gln	Lys	Ser	Tyr	Gly	Lys	Gly	Leu	Val	Gln	Thr	Thr	Arg	Gln	Leu	Pro
				340						345				350		
	Tyr	Asn	Gly	Val	Ile	Lys	Leu	Thr	Thr	Ala	Lys	Tyr	Tyr	Ile	Pro	Ser
				355				360					365			
30	Gly	Arg	Cys	Ile	Gln	Arg	Leu	Asp	Tyr	Ser	Arg	Thr	Asn	Arg	Thr	Gly
	370					375						380				
	Met	Ala	Thr	Ala	Ile	Pro	Asp	Ser	Leu	His	Lys	Ile	Phe	Tyr	Thr	Ala
	385					390						395				400
	Ala	Gly	Arg	Arg	Val	Glu	Asp	Ala	Gly	Gly	Ile	Leu	Pro	Asp	Ile	Glu
35				405						410				415		
	Val	Lys	Gln	Asp	Thr	Ala	Ala	Thr	Leu	Leu	Tyr	Tyr	Met	Ala	Ile	Asn
				420					425					430		
	Asn	Asp	Val	Phe	Asp	Phe	Val	Thr	Gly	Tyr	Val	Leu	Lys	His	Lys	Thr
				435					440					445		
40	Ile	Ala	Lys	Pro	Glu	Asp	Phe	Ser	Ile	Thr	Asn	Glu	Asp	Tyr	Ala	Ala

450 455 460
 Phe Cys Lys Met Met Glu Glu Lys Lys Phe Asp Tyr Asp Arg Gln Ser
 465 470 475 480
 Gly Lys Met Leu Asp Lys Leu Glu Glu Leu Ala Lys Ile Glu Gly Tyr
 5 485 490 495
 Leu Pro Glu Ala Asn Ser Glu Leu Lys Ala Leu Arg Glu Lys Leu Lys
 500 505 510
 Pro Asn Leu Ser Arg Asp Leu Leu Arg Phe Lys Lys Glu Ile Thr Asn
 515 520 525
 10 Tyr Leu Asn Asn Glu Ile Val Thr Arg Tyr Tyr Tyr Glu Arg Gly Ser
 530 535 540
 Ile Arg Gln Ser Leu Pro Glu Asp Lys Val Val Lys Glu Ala Ile Lys
 545 550 555 560
 Leu Leu Lys Asp His Pro Glu Gln Ile Arg Gln Ile Leu Ala Ala Pro
 15 565 570 575
 Lys Ala Glu Asn Lys Gly
 580

(2) INFORMATION FOR SEQ ID NO:17

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 295 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Porphyromonas gingivalis

(ix) FEATURE:

(A) NAME/KEY: misc_feature

35

(B) LOCATION 1...295

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

40 Thr Lys Lys Thr Leu Met Lys Lys Leu Phe Leu Ser Leu Thr Ser Leu
 1 5 10 15

	Val	Met	Val	Phe	Ala	Val	Ala	Ser	Cys	Asp	Ile	Ile	Asp	Lys	Asp	Gln
				20					25					30		
	Thr	Leu	Leu	Pro	Ala	Pro	Thr	Asn	Val	Thr	Pro	Asp	Asn	Pro	Asp	Asp
			35					40					45			
5	Asn	Pro	Ser	Glu	Ile	Asp	Ile	Thr	Gln	Thr	His	Thr	Glu	Lys	Tyr	Val
		50					55					60				
	Leu	Ala	Glu	Glu	Phe	Thr	Gly	Gln	Lys	Cys	Leu	Asn	Cys	Pro	Lys	Gly
	65					70					75				80	
	His	Arg	Lys	Leu	Ala	Ala	Leu	Lys	Glu	Gln	Tyr	Gly	Lys	Arg	Leu	Thr
10					85					90					95	
	Val	Val	Gly	Ile	His	Ala	Gly	Pro	Gly	Ser	Leu	Val	Pro	Pro	Leu	Phe
				100					105					110		
	Arg	Thr	Glu	Ala	Gly	Asp	Ala	Tyr	Tyr	Ser	Lys	Phe	Ala	Asn	Asn	Thr
			115					120					125			
15	Pro	Leu	Pro	Ala	Leu	Met	Val	Ser	Arg	Lys	Lys	Phe	Gly	Ser	Ser	Tyr
		130					135					140				
	Val	Tyr	Asp	Lys	Ser	Tyr	Lys	Thr	Trp	Asp	Val	Pro	Ile	Ala	Glu	Gln
	145					150					155				160	
	Met	Glu	Gln	Lys	Ala	Lys	Ile	Asn	Ile	Phe	Ala	Val	Ala	Glu	Tyr	Thr
20					165					170					175	
	Asp	Thr	Gln	Lys	Ile	Lys	Val	Thr	Val	Lys	Gly	Lys	Ile	Leu	Glu	Gly
			180						185					190		
	Asn	Thr	Leu	Pro	Lys	Ser	Met	Val	Gln	Val	Tyr	Leu	Leu	Glu	Asp	Lys
		195						200					205			
25	Leu	Ile	Ala	Pro	Gln	Val	Asp	Gly	Asn	Thr	Thr	Val	Glu	Asn	Tyr	Glu
		210					215					220				
	His	Asn	His	Val	Leu	Arg	Gly	Ala	Val	Asn	Gly	Ile	Trp	Gly	Glu	Glu
	225					230					235				240	
	Phe	Val	Asn	Leu	Lys	Asp	Tyr	Leu	Tyr	Thr	Tyr	Ala	Val	Glu	Pro	Leu
30					245					250					255	
	Ser	Gly	Met	Ser	Phe	Val	Ala	Glu	Asn	Tyr	Ser	Ile	Val	Ala	Phe	Val
			260						265					270		
	Tyr	Asp	Val	Gln	Thr	Phe	Glu	Val	Tyr	Asp	Val	Val	His	Val	Lys	Ile
			275					280					285			
35	Asn	Pro	Gln	Ser	Asp	Gly	Lys									
		290					295									

(2) INFORMATION FOR SEQ ID NO:18

40

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1046 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

10 (A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1046

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

	Asp	Phe	Pro	Trp	Val	Arg	Val	Lys	Pro	Glu	Lys	Lys	Arg	Lys	Gln	His	
	1				5					10					15		
20	Asn	Ser	Asn	Asn	Phe	Lys	Phe	Asn	Ala	Lys	Glu	Lys	Ser	Met	Lys	Arg	
				20					25					30			
	Met	Thr	Leu	Phe	Phe	Leu	Cys	Leu	Leu	Thr	Ser	Ile	Gly	Trp	Ala	Met	
			35				40						45				
	Ala	Gln	Asn	Arg	Thr	Val	Lys	Gly	Thr	Val	Ile	Ser	Ser	Glu	Asp	Asn	
25		50					55					60					
	Glu	Pro	Leu	Ile	Gly	Ala	Asn	Val	Val	Val	Val	Gly	Asn	Thr	Thr	Ile	
	65				70					75					80		
	Gly	Ala	Ala	Thr	Asp	Leu	Asp	Gly	Asn	Phe	Thr	Leu	Ser	Val	Pro	Ala	
				85					90					95			
30	Asn	Ala	Lys	Met	Leu	Arg	Val	Ser	Tyr	Ser	Gly	Met	Thr	Thr	Lys	Glu	
				100					105					110			
	Val	Ala	Ile	Ala	Asn	Val	Met	Lys	Ile	Val	Leu	Asp	Pro	Asp	Ser	Lys	
			115				120					125					
	Val	Leu	Glu	Gln	Val	Val	Val	Leu	Gly	Tyr	Gly	Thr	Gly	Gln	Lys	Leu	
35		130					135					140					
	Ser	Thr	Val	Ser	Gly	Ser	Val	Ala	Lys	Val	Ser	Ser	Glu	Lys	Leu	Ala	
	145				150				155					160			
	Glu	Lys	Pro	Val	Ala	Asn	Ile	Met	Asp	Ala	Leu	Gln	Gly	Gln	Val	Ala	
				165					170					175			
40	Gly	Met	Gln	Val	Met	Thr	Thr	Ser	Gly	Asp	Pro	Thr	Ala	Val	Ala	Ser	

	180	185	190
	Val Glu Ile His Gly Thr Gly Ser Leu Gly Ala Ser Ser Ala Pro Leu		
	195	200	205
	Tyr Ile Val Asp Gly Met Gln Thr Ser Leu Asp Val Val Ala Thr Met		
5	210	215	220
	Asn Pro Asn Asp Phe Glu Ser Met Ser Val Leu Lys Asp Ala Ser Ala		
	225	230	235
	Thr Ser Ile Tyr Gly Ala Arg Ala Ala Asn Gly Val Val Phe Ile Gln		
	245	250	255
10	Thr Lys Lys Gly Lys Met Ser Glu Arg Gly Arg Ile Thr Phe Asn Ala		
	260	265	270
	Ser Tyr Gly Ile Ser Gln Ile Leu Asn Thr Lys Pro Leu Asp Asn Met		
	275	280	285
	Met Thr Gly Asp Glu Leu Leu Asp Phe Gln Val Lys Ala Gly Phe Trp		
15	290	295	300
	Gly Asn Asn Gln Thr Val Gln Lys Val Lys Asp Met Ile Leu Ala Gly		
	305	310	315
	Ala Glu Asp Leu Tyr Gly Asn Tyr Asp Ser Leu Lys Asp Glu Tyr Gly		
	325	330	335
20	Lys Thr Leu Phe Pro Val Asp Phe Asn His Asp Ala Asp Trp Leu Lys		
	340	345	350
	Ala Leu Phe Lys Thr Ala Pro Thr Ser Gln Gly Asp Ile Ser Phe Ser		
	355	360	365
	Gly Gly Ser Gln Gly Thr Ser Tyr Tyr Ala Ser Ile Gly Tyr Phe Asp		
25	370	375	380
	Gln Glu Gly Met Ala Arg Glu Pro Ala Asn Phe Lys Arg Tyr Ser Gly		
	385	390	395
	Arg Leu Asn Phe Glu Ser Arg Ile Asn Glu Trp Leu Lys Val Gly Ala		
	405	410	415
30	Asn Leu Ser Gly Ala Ile Ala Asn Arg Arg Ser Ala Asp Tyr Phe Gly		
	420	425	430
	Lys Tyr Tyr Met Gly Ser Gly Thr Phe Gly Val Leu Thr Met Pro Arg		
	435	440	445
	Tyr Tyr Asn Pro Phe Asp Val Asn Gly Asp Leu Ala Asp Val Tyr Tyr		
35	450	455	460
	Met Tyr Gly Ala Thr Arg Pro Ser Met Thr Glu Pro Tyr Phe Ala Lys		
	465	470	475
	Met Arg Pro Phe Ser Ser Glu Ser His Gln Ala Asn Val Asn Gly Phe		
	485	490	495
40	Ala Gln Ile Thr Pro Ile Lys Gly Leu Thr Leu Lys Ala Gln Ala Gly		

	500	505	510
	Val Asp Ile Thr Asn Thr Arg Thr Ser Ser Lys Arg Met Pro Asn Asn		
	515	520	525
5	Pro Tyr Asp Ser Thr Pro Leu Gly Glu Arg Arg Glu Arg Ala Tyr Arg		
	530	535	540
	Asp Val Ser Lys Ser Phe Thr Asn Thr Ala Glu Tyr Lys Phe Ser Ile		
	545	550	555
	Asp Glu Lys His Asp Leu Thr Ala Leu Met Gly His Glu Tyr Ile Glu		
	565	570	575
10	Tyr Glu Gly Asp Val Ile Gly Ala Ser Ser Lys Gly Phe Glu Ser Asp		
	580	585	590
	Lys Leu Met Leu Leu Ser Gln Gly Lys Thr Gly Asn Ser Leu Ser Leu		
	595	600	605
15	Pro Glu His Arg Val Ala Glu Tyr Ala Tyr Leu Ser Phe Phe Ser Arg		
	610	615	620
	Phe Asn Tyr Gly Phe Asp Lys Trp Met Tyr Ile Asp Phe Ser Val Arg		
	625	630	635
	Asn Asp Gln Ser Ser Arg Phe Gly Ser Asn Asn Arg Ser Ala Trp Phe		
	645	650	655
20	Tyr Ser Val Gly Gly Met Phe Asp Ile Tyr Asn Lys Phe Ile Gln Glu		
	660	665	670
	Ser Asn Trp Leu Ser Asp Leu Arg Leu Lys Met Ser Tyr Gly Thr Thr		
	675	680	685
	Gly Asn Ser Glu Ile Gly Asn Tyr Asn His Gln Ala Leu Val Thr Val		
25	690	695	700
	Asn Asn Tyr Thr Glu Asp Ala Met Gly Leu Ser Ile Ser Thr Ala Gly		
	705	710	715
	Asn Pro Asp Leu Ser Trp Glu Lys Gln Ser Gln Phe Asn Phe Gly Leu		
	725	730	735
30	Ala Ala Gly Ala Phe Asn Asn Arg Leu Ser Ala Glu Val Asp Phe Tyr		
	740	745	750
	Val Arg Thr Thr Asn Asp Met Leu Ile Asp Val Pro Met Pro Tyr Ile		
	755	760	765
35	Ser Gly Phe Phe Ser Gln Tyr Gln Asn Val Gly Ser Met Lys Asn Thr		
	770	775	780
	Gly Val Asp Leu Ser Leu Lys Gly Thr Ile Tyr Gln Asn Lys Asp Trp		
	785	790	795
	Asn Val Tyr Ala Ser Ala Asn Phe Asn Tyr Asn Arg Gln Glu Ile Thr		
	805	810	815
40	Lys Leu Phe Phe Gly Leu Asn Lys Tyr Met Leu Pro Asn Thr Gly Thr		

	820	825	830
	Ile Trp Glu Ile Gly Tyr Pro Asn Ser Phe Tyr Met Ala Glu Tyr Ala		
	835	840	845
	Gly Ile Asp Lys Lys Thr Gly Lys Gln Leu Trp Tyr Val Pro Gly Gln		
5	850	855	860
	Val Asp Ala Asp Gly Asn Lys Val Thr Thr Ser Gln Tyr Ser Ala Asp		
	865	870	875
	Leu Glu Thr Arg Ile Asp Lys Ser Val Thr Pro Pro Ile Thr Gly Gly		880
	885	890	895
10	Phe Ser Leu Gly Ala Ser Trp Lys Gly Leu Ser Leu Asp Ala Asp Phe		
	900	905	910
	Ala Tyr Ile Val Gly Lys Trp Met Ile Asn Asn Asp Arg Tyr Phe Thr		
	915	920	925
	Glu Asn Unk Gly Gly Leu Met Gln Leu Asn Lys Asp Lys Met Leu Leu		
15	930	935	940
	Asn Ala Trp Thr Glu Asp Asn Lys Glu Thr Asp Val Pro Lys Leu Gly		
	945	950	955
	Gln Ser Pro Gln Phe Asp Thr His Leu Leu Glu Asn Ala Ser Phe Leu		
	965	970	975
20	Arg Leu Lys Asn Leu Lys Leu Thr Tyr Val Leu Pro Asn Ser Leu Phe		
	980	985	990
	Ala Gly Gln Asn Val Ile Gly Gly Ala Arg Val Tyr Leu Met Ala Arg		
	995	1000	1005
	Asn Leu Leu Thr Val Thr Lys Tyr Lys Gly Phe Asp Pro Glu Ala Gly		
25	1010	1015	1020
	Gly Asn Val Gly Lys Asn Gln Tyr Pro Asn Ser Lys Gln Tyr Val Ala		
	1025	1030	1035
	Gly Ile Gln Leu Ser Phe		1040
	1045		

30

(2) INFORMATION FOR SEQ ID NO:19

(i) SEQUENCE CHARACTERISTICS:

- 35 (A) LENGTH: 861 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

40 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

5 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...861

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19

10

```

Leu Glu Met Ala Tyr Asp Phe Thr Gln Thr Phe Arg Asn Ser Leu Glu
1           5           10           15
Tyr Ser Tyr Gln Glu Ala Thr Arg Leu Gly Val Val Ala Val Thr Gln
           20           25           30
15 Asp Met Leu Val Leu Gly Ile Ile Arg Asp Gly Asp Asn Gly Ala Ile
           35           40           45
Asp Ile Met Arg His Tyr Gly Ile Asn Leu Tyr Glu Leu Lys Arg Leu
           50           55           60
Ile Glu Leu Glu Ala Ile Ala Glu Ser Leu Pro Ala Ser Pro Glu Gly
20 65           70           75           80
Ser Pro Ile Phe Thr Pro Ser Ala Arg Glu Ala Ile Asp Asp Ala Thr
           85           90           95
Asp Ile Cys Ala Asp Met Glu Asp Glu Ala Val Ser Pro Val His Leu
           100          105          110
25 Leu Leu Ser Ile Leu Asn Ser Thr Gln Glu Ser Leu Val Gln Lys Ile
           115          120          125
Phe Met Lys Gln Gly Ile Lys Tyr Asp Thr Ile Leu Ser Asp Tyr Phe
           130          135          140
Gly Gln Arg Asn Pro Ser Glu Gly Lys Ser Pro Ser Glu Met Glu Ile
30 145          150          155          160
Leu Asp Gly Tyr Gln Asp Asn Asp Phe Asp Asp Glu Glu Asp Glu Ser
           165          170          175
Ser Pro Pro Ser Gly Asn Ser Gly Thr Gly Gly Gly Ser Gly Asp Ala
           180          185          190
35 Pro Glu Gln Asn Thr Gly Gly Gly Asp Thr Thr Thr Thr Thr Arg Ser
           195          200          205
Gly Gly Asp Thr Pro Ala Leu Asp Thr Phe Gly Thr Asp Ile Thr Ala
           210          215          220
Met Ala Ala Ala Gly Lys Leu Asp Pro Val Val Gly Arg Glu Gln Glu
40 225          230          235          240

```

	Ile	Glu	Arg	Val	Ile	Gln	Ile	Leu	Ser	Arg	Arg	Lys	Lys	Asn	Asn	Pro	
					245					250					255		
	Val	Leu	Ile	Gly	Glu	Pro	Gly	Val	Gly	Lys	Ser	Ala	Ile	Val	Glu	Gly	
				260					265					270			
5	Leu	Ala	Glu	Arg	Ile	Val	Asn	Arg	Lys	Val	Ser	Arg	Ile	Leu	Phe	Asp	
			275					280					285				
	Lys	Arg	Ile	Ile	Ser	Leu	Asp	Leu	Ala	Gln	Met	Val	Ala	Gly	Thr	Lys	
			290				295					300					
	Tyr	Arg	Gly	Gln	Phe	Glu	Glu	Arg	Leu	Lys	Ala	Val	Leu	Asp	Glu	Leu	
10	305					310					315					320	
	Lys	Lys	Asn	Pro	Gln	Ile	Ile	Leu	Phe	Ile	Asp	Glu	Ile	His	Thr	Ile	
					325					330					335		
	Val	Gly	Ala	Gly	Ser	Ala	Ala	Gly	Ser	Met	Asp	Thr	Ala	Asn	Met	Leu	
				340					345					350			
15	Lys	Pro	Ala	Leu	Ala	Arg	Gly	Gln	Val	Gln	Cys	Ile	Gly	Ala	Thr	Thr	
			355					360					365				
	Leu	Asp	Glu	Tyr	Arg	Lys	Asn	Ile	Glu	Lys	Asp	Gly	Ala	Leu	Glu	Arg	
			370				375					380					
	Arg	Phe	Gln	Lys	Val	Pro	Ile	Ala	Pro	Ser	Thr	Ala	Glu	Glu	Thr	Leu	
20	385					390					395					400	
	Thr	Ile	Leu	Gln	Asn	Ile	Lys	Glu	Lys	Tyr	Glu	Asp	Tyr	His	Gly	Val	
				405						410					415		
	Arg	Tyr	Thr	Asp	Glu	Ala	Ile	Lys	Ala	Ala	Val	Glu	Leu	Thr	Asp	Arg	
				420					425					430			
25	Tyr	Val	Ser	Asp	Arg	Phe	Phe	Pro	Asp	Lys	Ala	Ile	Asp	Ala	Met	Asp	
			435					440					445				
	Glu	Ala	Gly	Ala	Ser	Val	His	Ile	Thr	Asn	Val	Val	Ala	Pro	Lys	Glu	
			450				455					460					
	Ile	Glu	Ile	Leu	Glu	Ala	Glu	Leu	Ala	Ser	Val	Arg	Glu	Asn	Lys	Leu	
30	465					470					475					480	
	Ser	Ala	Val	Lys	Ala	Gln	Asn	Tyr	Glu	Leu	Ala	Ala	Ser	Phe	Arg	Asp	
				485						490					495		
	Gln	Glu	Arg	Arg	Thr	Gln	Gln	Gln	Ile	Ala	Glu	Glu	Lys	Lys	Lys	Trp	
				500					505					510			
35	Glu	Glu	Gln	Met	Ser	Lys	His	Arg	Glu	Thr	Val	Asp	Glu	Asn	Val	Val	
			515					520					525				
	Ala	His	Val	Val	Ala	Leu	Met	Thr	Gly	Val	Pro	Ala	Glu	Arg	Leu	Ser	
			530				535					540					
	Thr	Gly	Glu	Gly	Glu	Arg	Leu	Arg	Thr	Met	Ala	Asp	Asp	Leu	L		

Lys Val Val Gly Gln Asp Thr Ala Ile Glu Lys Met Val His Ala Ile
 565 570 575
 Gln Arg Asn Arg Leu Gly Leu Arg Asn Glu Lys Lys Pro Ile Gly Ser
 580 585 590
 5 Phe Leu Phe Leu Gly Pro Thr Gly Val Gly Lys Thr Tyr Leu Ala Lys
 595 600 605
 Lys Leu Ala Glu Tyr Leu Phe Glu Asp Glu Asn Ala Met Ile Arg Val
 610 615 620
 Asp Met Ser Glu Tyr Met Glu Lys Phe Ser Val Ser Arg Leu Val Gly
 10 625 630 635 640
 Ala Pro Pro Gly Tyr Val Gly Tyr Glu Glu Gly Gly Gln Leu Thr Glu
 645 650 655
 Arg Val Arg Arg Lys Pro Tyr Ser Val Val Leu Leu Asp Glu Ile Glu
 660 665 670
 15 Lys Ala His Ala Asp Val Phe Asn Leu Leu Leu Gln Val Met Asp Glu
 675 680 685
 Gly Gln Leu Thr Asp Ser Leu Gly Arg Arg Val Asn Phe Lys Asn Thr
 690 695 700
 Val Ile Ile Ile Thr Ser Asn Val Gly Thr Arg Gln Leu Lys Asp Phe
 20 705 710 715 720
 Gly Gln Gly Ile Gly Phe Arg Ser Glu Lys Asp Glu Glu Ala Asn Lys
 725 730 735
 Glu His Ser Arg Ser Val Ile Gln Lys Ala Leu Asn Lys Thr Phe Ser
 740 745 750
 25 Pro Glu Phe Leu Asn Arg Leu Asp Asp Ile Ile Leu Phe Asp Gln Leu
 755 760 765
 Gly Lys Thr Glu Ile Arg Arg Met Val Asp Ile Glu Leu Lys Ala Val
 770 775 780
 Leu Ala Arg Ile His Arg Ala Gly Tyr Asp Leu Val Leu Thr Asp Glu
 30 785 790 795 800
 Ala Lys Asp Val Ile Ala Thr Lys Gly Tyr Asp Leu Gln Tyr Gly Ala
 805 810 815
 Arg Pro Leu Lys Arg Thr Leu Gln Asn Glu Val Glu Asp Arg Leu Thr
 820 825 830
 35 Asp Leu Ile Leu Ser Gly Gln Ile Glu Lys Gly Gln Thr Leu Thr Leu
 835 840 845
 Ser Ala Arg Asp Gly Glu Ile Ile Val Gln Glu Gln Ala
 850 855 860

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 299 amino acids

(B) TYPE: amino acid

5

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

10

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

15

(A) NAME/KEY: misc_feature

(B) LOCATION 1...299

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

20 Leu Lys Lys Glu Ile Thr Met Lys Gln Asn Tyr Phe Lys Arg Val Cys
 1 5 10 15
 Ser Leu Leu Trp Leu Val Leu Pro Met Leu Ile Met Pro Leu Glu Val
 20 25 30
 Ala Ala Gln Glu Ile Ile Pro Asn Glu Glu Val Leu Glu Ser Leu Thr
 25 35 40 45
 Phe Val Ala Pro Val Glu Glu Thr Asp Ala Ile Glu Ala Glu Val Glu
 50 55 60
 Ala Leu Gln Glu Ile Val Ala Thr Glu Glu Ile Ala Glu Gln Ala Val
 65 70 75 80
 30 Arg Ser Tyr Thr Tyr Thr Val Tyr Arg Asp Gly Val Lys Ile Ala Ser
 85 90 95
 Gly Leu Thr Glu Pro Thr Phe Leu Asp Glu Asp Val Pro Ala Gly Glu
 100 105 110
 His Thr Tyr Cys Val Glu Val Gln Tyr Gln Gly Gly Val Ser Asp Lys
 35 115 120 125
 Val Cys Val Asp Val Glu Val Lys Asp Phe Lys Pro Val Thr Asn Leu
 130 135 140
 Thr Gly Thr Ala Ser Asn Asp Glu Val Ser Leu Asp Trp Asp Gly Val
 145 150 155 160
 40 Glu Glu Lys Ala Glu Glu Pro Ala Ser Asp Lys Ala Val Ser Tyr Asn

165 170 175
 Val Tyr Lys Asn Gly Thr Leu Ile Gly Asn Thr Ala Glu Thr His Tyr
 180 185 190
 Val Glu Thr Gly Val Ala Asn Gly Thr Tyr Ile Tyr Glu Val Glu Val
 5 195 200 205
 Lys Tyr Pro Asp Gly Val Ser Pro Lys Val Ala Val Thr Val Thr Val
 210 215 220
 Thr Asn Ser Ser Leu Ser Asn Val Asp Gly Gln Ala Pro Tyr Thr Leu
 225 230 235 240
 10 Arg Val Glu Gly Lys Lys Ile Ile Ala Glu Ala His Gly Met Ile Thr
 245 250 255
 Leu Tyr Asp Ile Asn Gly Arg Thr Val Ala Val Ala Pro Asn Arg Leu
 260 265 270
 Glu Tyr Met Ala Gln Thr Gly Phe Tyr Ala Val Arg Phe Asp Val Gly
 15 275 280 285
 Asn Lys His His Val Ser Lys Ile Gln Val Arg
 290 295

(2) INFORMATION FOR SEQ ID NO:21

20

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 464 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Porphyromonas gingivalis

(ix) FEATURE:

(A) NAME/KEY: misc_feature

35

(B) LOCATION 1...464

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

Tyr Gly Lys Arg Arg Lys Leu Gly Thr Ser Val Arg Pro Ser Val Leu
 40 1 5 10 15

Thr Gln Ile Arg Phe Ile Leu Asp Leu His Leu Ile Thr Asp Phe Phe
 20 25 30
 Glu Gly Leu Arg Val Asn Pro Ile Gly Ala Ala Ala Ile Val Ala Phe
 35 40 45
 5 Ile Ile Asp Leu Leu Leu Leu Cys Cys Ser Ala Phe Met Ser Ser Cys
 50 55 60
 Glu Val Ala Tyr Phe Ser Leu Lys Pro Ile Asp Leu Gln Asn Ile Arg
 65 70 75 80
 Glu Arg Asn His Ser Ser Asp Ile Ala Leu Ser Asn Leu Leu Asp Asn
 10 85 90 95
 Ser Asn Gln Leu Leu Ala Thr Ile Leu Ile Gly Asn Asn Val Ile Asn
 100 105 110
 Val Ala Ile Val Ile Leu Ser Asn Tyr Ala Ile Glu Gln Thr Phe Val
 115 120 125
 15 Phe Ser Ser Pro Ile Ile Gly Phe Leu Ile Gln Thr Ile Leu Leu Thr
 130 135 140
 Thr Val Leu Leu Leu Phe Gly Glu Ile Leu Pro Lys Val Tyr Ala Arg
 145 150 155 160
 Lys Asn Pro Leu Gln Tyr Ser Arg Phe Ser Ala Ala Ala Met Ser Val
 165 170 175
 20 Ile Tyr Lys Ile Leu Ser Pro Phe Ser Lys Leu Leu Val Lys Ser Thr
 180 185 190
 Gly Ile Val Thr Arg Gly Ile Ser Lys Lys Lys Tyr Asp Met Ser Val
 195 200 205
 25 Asp Glu Leu Ser Lys Ala Val Ala Leu Thr Thr Thr Glu Gly Glu Pro
 210 215 220
 Glu Glu Lys Glu Met Ile Asn Glu Ile Ile Lys Phe Tyr Asn Lys Thr
 225 230 235 240
 Ala Cys Glu Ile Met Val Pro Arg Ile Asp Ile Val Asp Val Asp Leu
 245 250 255
 30 Ser Trp Pro Phe Arg Lys Met Leu Asp Phe Val Val Ser Ser Gly Tyr
 260 265 270
 Ser Arg Leu Pro Val Ser Glu Gly Ser Glu Asp Asn Ile Lys Gly Val
 275 280 285
 35 Ile Tyr Ile Lys Asp Leu Ile Pro His Met Asp Lys Gly Asp Glu Phe
 290 295 300
 Asp Trp His Pro Leu Ile Arg Lys Ala Tyr Phe Val Pro Glu Asn Lys
 305 310 315 320
 Arg Ile Asp Asp Leu Leu Glu Glu Phe Arg Ala Asn Lys Val His Val
 325 330 335
 40

```

Ser Ile Val Val Asp Glu Phe Gly Gly Thr Cys Gly Leu Ile Thr Met
      340                      345                      350
Glu Asp Ile Leu Glu Glu Ile Val Gly Glu Ile Thr Asp Glu Tyr Asp
      355                      360                      365
5  Glu Glu Glu Leu Pro Phe Lys Val Leu Gly Asp Gly Ser Tyr Leu Phe
      370                      375                      380
Glu Gly Lys Thr Ser Leu Ser Asp Val Arg His Tyr Leu Asp Leu Pro
385                      390                      395                      400
Glu Asn Ala Phe Gly Glu Leu Gly Asp Glu Val Asp Thr Leu Ser Gly
10                      405                      410                      415
Leu Phe Leu Glu Ile Lys Gln Glu Leu Pro His Val Gly Asp Thr Ala
      420                      425                      430
Val Tyr Glu Pro Phe Arg Phe Gln Val Thr Gln Met Asp Lys Arg Arg
      435                      440                      445
15  Ile Ile Glu Ile Lys Ile Phe Pro Phe Glu Arg Thr Trp Glu Val Glu
      450                      455                      460

```

(2) INFORMATION FOR SEQ ID NO:22

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 869 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
- 30 (A) ORGANISM: *Porphyromonas gingivalis*
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...869
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22

```

Trp His Arg Asn Ile Phe Ile Phe Ala Ser Thr Phe Ser Pro Lys Asn
1          5          10          15
40 Met Leu Pro Leu Pro Tyr Arg Tyr Ala Lys Thr Glu His Leu Phe Leu

```

	20	25	30
	Ala Lys Gly Tyr Cys Lys Asn Pro Ile Thr Asn Ile Ile Ile Leu Phe		
	35	40	45
	Met Lys Lys Lys Asn Phe Leu Leu Leu Gly Ile Phe Val Ala Leu Leu		
5	50	55	60
	Thr Phe Ile Gly Ser Met Gln Ala Gln Gln Ala Lys Asp Tyr Phe Asn		
	65	70	75
	Phe Asp Glu Arg Gly Glu Ala Tyr Phe Ser Phe Lys Val Pro Asp Arg		
	85	90	95
10	Ala Val Leu Gln Glu Leu Ala Leu Ile Met Ser Ile Asp Glu Phe Asp		
	100	105	110
	Pro Val Thr Asn Glu Ala Ile Ala Tyr Ala Ser Glu Glu Glu Phe Glu		
	115	120	125
	Ala Phe Leu Arg Tyr Gly Leu Lys Pro Thr Phe Leu Thr Pro Pro Ser		
15	130	135	140
	Met Gln Arg Ala Val Glu Met Phe Asp Tyr Arg Ser Gly Glu Lys Tyr		
	145	150	155
	Glu Trp Asn Ala Tyr Pro Thr Tyr Glu Ala Tyr Ile Ser Met Met Glu		
	165	170	175
20	Glu Phe Gln Thr Lys Tyr Pro Ser Leu Cys Thr Thr Ser Val Ile Gly		
	180	185	190
	Lys Ser Val Lys Asp Arg Lys Leu Met Ile Cys Lys Leu Thr Ser Ser		
	195	200	205
	Ala Asn Thr Gly Lys Lys Pro Arg Val Leu Tyr Thr Ser Thr Met His		
25	210	215	220
	Gly Asp Glu Thr Thr Gly Tyr Val Val Leu Leu Arg Leu Ile Asp His		
	225	230	235
	Leu Leu Ser Asn Tyr Glu Ser Asp Pro Arg Ile Lys Asn Ile Leu Asp		
	245	250	255
30	Lys Thr Glu Val Trp Ile Cys Pro Leu Thr Asn Pro Asp Gly Ala Tyr		
	260	265	270
	Arg Ala Gly Asn His Thr Val Gln Gly Ala Thr Arg Tyr Asn Ala Asn		
	275	280	285
	Asn Val Asp Leu Asn Arg Asn Phe Lys Asp Asp Val Ala Gly Asp His		
35	290	295	300
	Pro Asp Gly Lys Pro Trp Gln Pro Glu Ala Thr Ala Phe Met Asp Leu		
	305	310	315
	Glu Gly Asn Thr Ser Phe Val Leu Gly Ala Asn Ile His Gly Gly Thr		
	325	330	335
40	Glu Val Val Asn Tyr Pro Trp Asp Asn Lys Lys Glu Arg His Ala Asp		

	340	345	350
	Asp Glu Trp Tyr Lys Leu Ile Ser Arg Asn Tyr Ala Ala Ala Cys Gln		
	355	360	365
	Ser Ile Ser Ala Ser Tyr Met Thr Ser Glu Thr Asn Ser Gly Ile Ile		
5	370	375	380
	Asn Gly Ser Asp Trp Tyr Val Ile Arg Gly Ser Arg Gln Asp Asn Ala		
	385	390	395
	Asn Tyr Phe His Arg Leu Arg Glu Ile Thr Leu Glu Ile Ser Asn Thr		
	405	410	415
10	Lys Leu Val Pro Ala Ser Gln Leu Pro Lys Tyr Trp Asn Leu Asn Lys		
	420	425	430
	Glu Ser Leu Leu Ala Leu Ile Glu Glu Ser Leu Tyr Gly Ile His Gly		
	435	440	445
	Thr Val Thr Ser Ala Ala Asn Gly Gln Pro Leu Lys Cys Gln Ile Leu		
15	450	455	460
	Ile Glu Asn His Asp Lys Arg Asn Ser Asp Val Tyr Ser Asp Ala Thr		
	465	470	475
	Thr Gly Tyr Tyr Val Arg Pro Ile Lys Ala Gly Thr Tyr Thr Val Lys		
	485	490	495
20	Tyr Lys Ala Glu Gly Tyr Pro Glu Ala Thr Arg Thr Ile Thr Ile Lys		
	500	505	510
	Asp Lys Glu Thr Val Ile Met Asp Ile Ala Leu Gly Asn Ser Val Pro		
	515	520	525
	Leu Pro Val Pro Asp Phe Thr Ala Ser Pro Met Thr Ile Ser Val Gly		
25	530	535	540
	Glu Ser Val Gln Phe Gln Asp Gln Thr Thr Asn Asn Pro Thr Asn Trp		
	545	550	555
	Glu Trp Thr Phe Glu Gly Gly Gln Pro Ala Met Ser Thr Glu Gln Asn		
	565	570	575
30	Pro Leu Val Ser Tyr Ser His Pro Gly Gln Tyr Asp Val Thr Leu Lys		
	580	585	590
	Val Trp Asn Ala Ser Gly Ser Asn Thr Ile Thr Lys Glu Lys Phe Ile		
	595	600	605
	Thr Val Asn Ala Val Met Pro Val Ala Glu Phe Val Gly Thr Pro Thr		
35	610	615	620
	Glu Ile Glu Glu Gly Gln Thr Val Ser Phe Gln Asn Gln Ser Thr Asn		
	625	630	635
	Ala Thr Asn Tyr Val Trp Ile Phe Asp Gly Gly Thr Pro Ala Thr Ser		
	645	650	655
40	Glu Asp Glu Asn Pro Thr Val Leu Tyr Ser Lys Ala Gly Gln Tyr Asp		

[illegible]

(2) INFORMATION FOR SEQ ID NO:23

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 843 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

40

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc_feature

5 (B) LOCATION 1...843

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

```

10  Met Lys Lys Ser Phe Leu Leu Ala Ile Val Met Leu Phe Gly Ile Ala
    1           5           10           15
    Met Gln Gly His Ser Ala Pro Val Thr Lys Glu Arg Ala Leu Ser Leu
        20           25           30
    Ala Arg Leu Ala Leu Arg Gln Val Ser Leu Arg Met Gly Gln Thr Ala
        35           40           45
15  Val Ser Asp Lys Ile Ser Ile Asp Tyr Val Tyr Arg Gln Gly Asp Ala
    50           55           60
    Glu Arg Gly Ile Thr Ser Gln Glu Glu Gly Ser Pro Ala Tyr Phe Tyr
    65           70           75           80
    Val Ala Asn Arg Gly Asn Asn Glu Gly Tyr Ala Leu Val Ala Ala Asp
20           85           90           95
    Asp Arg Ile Pro Thr Ile Leu Ala Tyr Ser Pro Ile Gly Arg Phe Asp
        100           105           110
    Met Asp Ser Met Pro Asp Asn Leu Arg Met Trp Leu Gln Ile Tyr Asp
        115           120           125
25  Gln Glu Ile Gly Leu Ile Leu Ser Gly Lys Ala Gln Leu Asn Glu Glu
    130           135           140
    Ile Leu Arg Thr Glu Gly Val Pro Ala Glu Val His Ala Leu Met Asp
    145           150           155           160
    Asn Gly His Phe Ala Asn Asp Pro Met Arg Trp Asn Gln Gly Tyr Pro
30           165           170           175
    Trp Asn Asn Lys Glu Pro Leu Leu Pro Asn Gly Asn His Ala Tyr Thr
        180           185           190
    Gly Cys Val Ala Thr Ala Ala Ala Gln Ile Met Arg Tyr His Ser Trp
        195           200           205
35  Pro Leu Gln Gly Glu Gly Ser Phe Asp Tyr His Ala Gly Ser Leu Val
    210           215           220
    Gly Asn Trp Ser Gly Thr Phe Gly Glu Met Tyr Asp Trp Ile Asn Met
    225           230           235           240
    Pro Gly Asn Pro Asp Leu Asp Asn Leu Thr Gln Ser Gln Val Asp Ala
40           245           250           255

```

Tyr Ala Thr Leu Met Arg Asp Val Ser Ala Ser Val Ser Met Ser Phe
 260 265 270
 Tyr Glu Asn Gly Ser Gly Thr Tyr Ser Val Tyr Val Val Gly Ala Leu
 275 280 285
 5 Arg Asn Asn Phe Arg Tyr Lys Arg Ser Leu Gln Leu His Val Arg Ala
 290 295 300
 Leu Tyr Thr Ser Gln Glu Trp His Asp Met Ile Arg Gly Glu Leu Ala
 305 310 315 320
 Ser Gly Arg Pro Val Tyr Tyr Ala Gly Asn Asn Gln Ser Ile Gly His
 10 325 330 335
 Ala Phe Val Cys Asp Gly Tyr Ala Ser Asp Gly Thr Phe His Phe Asn
 340 345 350
 Trp Gly Trp Gly Gly Val Ser Asn Gly Phe Tyr Lys Leu Thr Leu Leu
 355 360 365
 15 Ser Pro Thr Ser Leu Gly Ile Gly Gly Glu Gly Ile Gly Phe Thr Ile
 370 375 380
 Tyr Gln Glu Ile Ile Thr Gly Ile Glu Pro Ala Lys Thr Pro Ala Glu
 385 390 395 400
 Ala Gly Thr Asp Ala Leu Pro Ile Leu Ala Leu Lys Asp Ile Glu Ala
 20 405 410 415
 Glu Tyr Lys Ser Glu Ser Gly Leu Asn Val Gly Tyr Ser Ile Tyr Asn
 420 425 430
 Thr Gly Glu Glu Gln Ser Asn Leu Asp Leu Gly Tyr Arg Leu Asn Lys
 435 440 445
 25 Ala Asp Gly Glu Val Ile Glu Val Lys Thr Ser Ser Ile Asn Ile Ser
 450 455 460
 Trp Tyr Gly Tyr Gly Glu His Pro Glu Ser Phe Ser Leu Ala Pro Asn
 465 470 475 480
 Gln Leu Ser Gln Gly Ile Asn Thr Ile Thr Leu Leu Tyr Arg Arg Thr
 30 485 490 495
 Gly Thr Glu Gln Trp Glu Pro Val Arg His Ala Gln Gly Gly Tyr Val
 500 505 510
 Asn Ser Ile Lys Val Asn Thr Thr Asp Pro Asn Asn Val Val Val Thr
 515 520 525
 35 Val Asp Asn Asn Glu Gly Lys Leu Ser Ile Val Pro Asn Ser Phe Val
 530 535 540
 Ala Asp Leu Asn Ser Tyr Glu His Ser Thr Ile Thr Val Gln Phe Asn
 545 550 555 560
 Ser Asp Ser Pro Asp Glu Ile Arg Thr Pro Val Ala Phe Ala Leu Ser
 40 565 570 575

Thr Gly Ala Thr Ala Asp Asp Val Ile Ser Leu Gly Trp Val Met Ala
 580 585 590
 Glu Val Pro Gly Gly Ser Ser Asn Tyr Pro Val Val Trp Ser Lys Asp
 595 600 605
 5 Val Leu Thr Leu Ser Glu Gly Asp Tyr Thr Leu Trp Tyr Arg Phe Ser
 610 615 620
 Ile Asn Asn Gln Lys Asp Glu Trp Lys Lys Ile Gly Ser Val Ser Val
 625 630 635 640
 Lys Thr Pro Thr Glu Tyr Thr His Pro Leu Phe Glu Val Gly His Asn
 10 645 650 655
 Gln Thr Ser Thr Tyr Thr Leu Asp Met Ala His Asn Arg Val Leu Pro
 660 665 670
 Asp Phe Thr Leu Lys Asn Leu Gly Leu Pro Phe Asn Gly Glu Leu Val
 675 680 685
 15 Val Val Phe Arg Gln Thr Gln Ser Ser Ser Gly Ser Leu Trp Ala Ala
 690 695 700
 Gln Glu Thr Val His Ile Lys Gln Gly Glu Thr Phe Val Tyr Lys Pro
 705 710 715 720
 Val Val Glu Gly Pro Ile Pro Asp Gly Ser Tyr Arg Ala Thr Leu His
 20 725 730 735
 Ala Phe Val Asn Gly Gln Gln Gln Leu Tyr Leu Lys Gly Lys Arg Asn
 740 745 750
 Tyr Thr Val Lys Ile Val Asn Gly Thr Ala Val Glu Ala Ile Glu Ser
 755 760 765
 25 Ser Glu Glu Ile Arg Val Phe Pro Asn Pro Ala Arg Asp Tyr Val Glu
 770 775 780
 Ile Ser Ala Pro Cys Ile Pro Gln Glu Thr Ser Ile Ile Leu Phe Asp
 785 790 795 800
 Leu Ser Gly Lys Ile Val Met Lys Asn Ser Leu Ser Ala Gly His Gly
 30 805 810 815
 Arg Met Asp Val Ser Arg Leu Pro Asn Gly Ala Tyr Ile Leu Lys Val
 820 825 830
 Asp Gly Tyr Thr Thr Lys Ile Asn Ile Val His
 835 840
 35

(2) INFORMATION FOR SEQ ID NO:24

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1175 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

10 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...1175

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

15

Met Thr Trp Leu Ile Asp Phe Pro Ala Leu Lys Lys Leu Asp Leu Ser
1 5 10 15

Tyr Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu
20 25 30

Thr Lys Leu Arg Leu Arg Ser Asn Gln Ile Arg Lys Leu Glu Gly Leu
35 40 45

Asp Ser Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile
50 55 60

Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr
25 65 70 75 80

Leu Leu Asp Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr
85 90 95

Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys Leu Glu
100 105 110

Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn
30 115 120 125

Gln Ile Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Thr Lys
130 135 140

Leu Arg Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg
35 145 150 155 160

Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser Gly Asn Gln Ile Arg Lys
165 170 175

Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Thr Leu Glu Leu Ser
180 185 190

40 Gly Asn Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Ser Ser Leu

	195	200	205
	Thr Lys Leu Arg Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu Gly Leu		
	210	215	220
	Glu Arg Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn Gln Ile		
5	225	230	235 240
	Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu Leu Tyr		
	245	250	255
	Leu Leu Asp Asn Gln Ile Arg Lys Leu Glu Gly Leu Glu Arg Leu Thr		
	260	265	270
10	Ser Leu Thr Lys Leu Arg Leu Arg Ser Asn Gln Ile Ser Lys Leu Glu		
	275	280	285
	Gly Leu Asp Ser Leu Thr Ser Leu Thr Lys Leu Ser Leu Ser Asp Asn		
	290	295	300
	Gln Ile Ser Lys Leu Glu Gly Leu Glu Arg Leu Thr Ser Leu Ala Glu		
15	305	310	315 320
	Leu Tyr Leu Leu Asp Asn Gln Ile Arg Lys Leu Glu Gly Leu Asp Gly		
	325	330	335
	Leu Ala Ser Leu Thr Arg Leu Ser Leu Arg Arg Asn Gln Ile Ser Lys		
	340	345	350
20	Leu Glu Gly Leu Asp Arg Leu Lys Val Leu Arg Lys Leu Asp Val Ser		
	355	360	365
	Gly Asn Asp Ile Gln Ser Ile Asp Asp Ile Lys Leu Leu Ala Pro Ile		
	370	375	380
	Leu Glu Gln Thr Leu Glu Lys Leu Arg Ile His Asp Asn Pro Phe Val		
25	385	390	395 400
	Ala Ser Ser Gly Leu Ile Leu Ser Pro Tyr Asp Asn His Leu Pro Glu		
	405	410	415
	Ile Lys Ala Leu Leu Glu Lys Glu Lys Glu Lys Gln Lys Lys Thr Ser		
	420	425	430
30	Val Glu Tyr His Pro Phe Cys Lys Val Met Leu Leu Gly Asn His Ser		
	435	440	445
	Ser Gly Lys Thr Thr Phe Leu Ser Gln Tyr Asp Thr Asn Tyr Thr Tyr		
	450	455	460
	Gln Lys Asn Thr His Val Leu Ser Ile His Arg Ser Asn Asn Pro Asn		
35	465	470	475 480
	Ala Ile Phe Tyr Asp Phe Gly Gly Gln Asp Tyr Tyr His Gly Ile Tyr		
	485	490	495
	Gln Ala Phe Phe Thr Thr Gln Ser Leu Tyr Leu Leu Phe Trp Asp Ala		
	500	505	510
40	Lys Lys Asp Arg Asn Phe Val Ser Val Asp Asp Lys Glu Tyr Gln Thr		

	515		520		525
	Leu Asn Phe Asn Arg Pro Tyr Trp	Leu Gly Gln Ile Ala Tyr Ala Cys			
	530	535	540		
5	Asn Arg Cys Met Ser Val Gly Gly Asn Pro Asp Gly Lys Asp Thr Pro				
	545	550	555	560	
	Gln Thr Thr Asp Asp Thr Ile Ile Ile Gln Thr His Ala Asp Glu Thr				
	565	570	575		
	Gly Ala Lys Gln Gln Thr Leu Gly Cys Ala Ala Glu Asn Gly Val Leu				
	580	585	590		
10	Glu Glu Ile Tyr Val Ser Leu Glu Pro Lys Ala Asn Ser Ala Val His				
	595	600	605		
	Ala Leu Asn Tyr Leu Asn Glu Arg Val Arg Glu Val Val Ala Ser Arg				
	610	615	620		
	Ser Lys Ser Ile Gln Ile Thr Glu Lys Asp Lys Gly Leu Tyr Glu Ala				
15	625	630	635	640	
	Leu Pro Thr Ile Ala Gly Asp Asn Lys His Ile Pro Ile Ser Leu Glu				
	645	650	655		
	Ala Leu Ala Ala Gln Leu Asn Lys Gly Arg Ala Glu Asn Asp Leu Tyr				
	660	665	670		
20	Thr Ile Glu Tyr Leu Gln Thr Glu Leu Asn Gln Leu Ser Leu Arg Gly				
	675	680	685		
	Glu Val Leu Tyr Tyr Arg Glu Asn Glu Lys Leu Asn Asn Tyr Val Trp				
	690	695	700		
	Leu Asp Pro Ala Ala Phe Val Gln Met Ile His Gly Glu Ile Leu Gln				
25	705	710	715	720	
	Lys Asp Asn Ile Asn Arg Gly Thr Val Pro Lys Asp Ile Phe Glu Cys				
	725	730	735		
	Lys Leu His Asn Leu Ser Ser Gly Ser Ile Phe Glu Glu Asp Gly Gln				
	740	745	750		
30	Asn Gly Asn Met Ile Leu Gln Leu Leu Leu Glu Glu Leu Ile Val Tyr				
	755	760	765		
	Glu Asp Lys Asp Cys Tyr Val Ile Pro Gly Tyr Leu Pro Leu His Ser				
	770	775	780		
	Asp Asp Glu Ala Tyr Lys Trp Leu Thr Leu Gly Phe Glu Arg Pro Asn				
35	785	790	795	800	
	Phe Val Leu Lys Phe Glu Arg Phe Ile Pro Phe Gly Leu Ile Asn Gln				
	805	810	815		
	Ile Ile Ala Tyr Tyr Gly Arg Glu Glu Gly Ala Leu Lys Arg Tyr Trp				
	820	825	830		
40	Arg Asp Gln Val Ile Phe Thr Ala Gly Arg Glu Met Asp Arg Gln Thr				

	835		840		845
	Leu Glu Gln Glu Glu Glu Lys Glu Gly Leu Pro Lys Thr Asn Ala Glu				
	850		855		860
5	Asp Tyr Gln Ile Trp Ile Lys Leu Asp Phe Thr Asp Leu Ala Ile Ser				
	865		870		875
	Val Phe Ile Lys Glu Gln Arg Lys Thr Ser Ala Lys Asp Met Gln Arg				
		885		890	895
	Lys Glu Ala Thr Ile Leu Ser Asp Met Leu Asp Met Tyr Trp Asn Asn				
		900		905	910
10	Ile Pro Pro Arg Glu Gln Ile Gly Asp Lys Asp Thr Glu Gln Thr Arg				
		915		920	925
	Ser Thr Ile Arg Glu Thr Asn Arg Lys Lys Arg Pro Ile Gln Asp Leu				
		930		935	940
15	Tyr Leu Ser Cys Ala Gln Ala Asp Lys Asp Leu Thr Glu Ser His Tyr				
	945		950		955
	Ile His Leu Gly Thr Leu Asp Asp Glu Ser Lys Thr Thr Ala Arg Ile				
		965		970	975
	Ala Ala Tyr Pro Leu Lys Asn Gly Val Ile Asp Lys Glu Arg Val Arg				
		980		985	990
20	Glu Val Ser Thr Arg Pro Tyr Lys His Leu Ser Val Asn Lys Asn Leu				
		995		1000	1005
	Ala Thr Ala Lys Gln Ile Phe Ile Ser Tyr Ser Lys Glu Asp Gln Thr				
		1010		1015	1020
25	Glu Leu Glu Thr Cys Leu Gln Phe Phe Lys Pro Leu Glu Lys Asn Gly				
	1025		1030		1035
	Gln Ile Glu Ile Tyr Tyr Asp Lys Leu Thr Lys Phe Glu Thr Pro Ile				
		1045		1050	1055
	His Pro Glu Ile Arg Lys Arg Ile Val Glu Ala Asp Cys Ile Ile Ala				
		1060		1065	1070
30	Leu Ile Ser Gln Arg Tyr Leu Ala Thr Asp Tyr Ile Leu Asp His Glu				
		1075		1080	1085
	Leu Pro Val Phe Arg Glu Tyr Asn Lys Thr Ile Val Pro Ile Leu Ile				
		1090		1095	1100
35	Lys Pro Cys Thr Phe Glu Asp Asp Glu Phe Leu Arg Glu Lys Tyr Phe				
	1105		1110		1115
	Ala Gln Lys Ala Gln Ile Ile Asn Leu Gly Lys Glu Gly Lys Thr Ile				
		1125		1130	1135
	Lys Ala Tyr Asp Ser Ile Thr Ala Ser Ala His Arg Asp Glu Asn Trp				
		1140		1145	1150
40	Val Ala Val Val Arg Glu Phe Lys Glu Lys Ile Leu Arg Ile Thr Lys				

1155 1160 1165
 Gln Glu Val Asn Thr Asp Glu
 1170 1175

5 (2) INFORMATION FOR SEQ ID NO:25

(i) SEQUENCE CHARACTERISTICS:

- 10 (A) LENGTH: 312 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

15 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Porphyromonas gingivalis

20 (ix) FEATURE:

- (A) NAME/KEY: misc_feature
 (B) LOCATION 1...312

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25

25 Met Ile Pro Leu Ser Glu Ser Phe Glu Ser Gly Ile Pro Ala Ile Trp
 1 5 10 15
 Lys Thr Ile Asp Ala Asp Gly Asp Gly Tyr Asn Trp Met His Leu Thr
 20 25 30
 Asn Phe Thr Gly Gln Ser Gly Leu Cys Val Ser Ser Ala Ser Tyr Ile
 30 35 40 45
 Gly Gly Val Gly Ala Leu Thr Pro Asp Asn Tyr Leu Ile Thr Pro Glu
 50 55 60
 Leu Lys Leu Pro Thr Asp Ala Leu Val Glu Ile Ile Tyr Trp Val Cys
 65 70 75 80
 35 Thr Gln Asp Leu Thr Ala Pro Ser Glu His Tyr Ala Val Tyr Ser Ser
 85 90 95
 Ser Thr Gly Asn Asn Ala Ala Asp Phe Val Asn Leu Leu Tyr Glu Glu
 100 105 110
 Thr Leu Thr Ala Lys Arg Ile Gln Ser Pro Glu Leu Ile Arg Gly Asn
 40 115 120 125

Arg Thr Gln Gly Val Trp Tyr Gln Arg Lys Val Val Leu Pro Asn Asp
 130 135 140
 Thr Lys Tyr Val Ala Phe Arg His Phe Asn Ser Thr Asp Asn Phe Trp
 145 150 155 160
 5 Leu Asn Leu Asp Glu Val Ser Ile Leu Tyr Thr Pro Leu Pro Arg Arg
 165 170 175
 Ala Pro Cys Pro His Pro Gly Gly Tyr Thr Tyr Ser Val Phe Arg Asp
 180 185 190
 Gly Gln Lys Ile Ala Ser Gly Leu Ser Ala Leu Ala Tyr Ile Asp Thr
 10 195 200 205
 Asp Val Pro Tyr Gly Thr Gln Asp Tyr Cys Val Gln Val Asn Tyr Leu
 210 215 220
 Gln Gly Asp Ser Tyr Lys Val Cys Lys Asn Ile Val Val Ala Asn Ser
 225 230 235 240
 15 Ala Asn Ile Tyr Gly Ala Asp Lys Pro Phe Ala Leu Thr Val Val Gly
 245 250 255
 Lys Thr Ile Val Ala Ser Ala Phe Lys Gly Glu Ile Thr Leu Tyr Asp
 260 265 270
 Ile Arg Gly Arg Leu Ile Ala Ser Gly Cys Asp Thr Leu Arg Tyr Lys
 20 275 280 285
 Ala Glu Asn Gly Phe Tyr Leu Ile Lys Ile Gln Val Asn Gly Thr Val
 290 295 300
 Tyr Thr Glu Lys Ile Gln Ile Gln
 305 310

(2) INFORMATION FOR SEQ ID NO:26

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 757 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

40 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...757

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26

5

```

Met Arg Thr Lys Thr Ile Phe Phe Ala Ile Ile Ser Phe Ile Ala Leu
1           5           10           15
Leu Ser Ser Ser Leu Ser Ala Gln Ser Lys Ala Val Leu Thr Gly Ser
           20           25           30
10 Val Ser Asp Ala Glu Thr Gly Glu Pro Leu Ala Gly Ala Arg Ile Glu
           35           40           45
Val Lys His Thr Asn Ile Val Ala Gly Ala Asp Ala Gly Gly His Phe
           50           55           60
Glu Ile Lys Asn Leu Pro Ala Gly Gln His Thr Ile Ile Cys Ser Leu
15 65           70           75           80
Gly Gly Tyr Gly Gln Lys Glu Glu Val Val Ala Ile Glu Ala Gly Gln
           85           90           95
Thr Lys Thr Ile Ser Phe Ala Leu Arg Leu Arg Thr Asn Asn Leu Glu
           100          105          110
20 Glu Val Val Val Thr Gly Thr Gly Thr Arg Tyr Arg Leu Val Asp Ala
           115          120          125
Pro Val Ala Thr Glu Val Leu Thr Ala Lys Asp Ile Ala Ser Phe Ser
           130          135          140
Ala Pro Thr Ser Glu Ala Leu Leu Gln Gly Leu Ser Pro Ser Phe Asp
25 145          150          155          160
Phe Gly Pro Asn Leu Met Gly Ser Phe Met Gln Leu Asn Gly Leu Ser
           165          170          175
Ser Lys Tyr Ile Leu Ile Leu Ile Asp Gly Lys Arg Val Tyr Gly Asp
           180          185          190
30 Val Gly Gly Gln Ala Asp Leu Ser Arg Ile Ser Pro Asp Gln Ile Glu
           195          200          205
Arg Ile Glu Leu Val Lys Gly Ala Ser Ser Ser Leu Tyr Gly Ser Asp
           210          215          220
Ala Ile Ala Gly Val Ile Asn Val Ile Thr Lys Lys Asn Thr Asn Arg
35 225          230          235          240
Leu Ser Ala Tyr Thr Ser His Arg Ile Ser Lys Tyr Asn Asp Arg Gln
           245          250          255
Thr Asn Thr Ser Leu Asp Ile Asn Ile Gly Lys Phe Ser Ser Asn Thr
           260          265          270
40 Asn Tyr Phe Phe Tyr His Thr Asp Gly Trp Gln Asn Ser Pro Phe Glu

```


	595		600		605											
	Tyr	Ala	Asn	Ile	Gly	Lys	Ala	Arg	Ser	Arg	Gly	Leu	Asp	Val	Leu	Cys
	610						615					620				
	Asp	Ala	Ser	Ile	Gly	Trp	Gly	Ile	Lys	Leu	Gly	Ala	Gly	Tyr	Ser	Leu
5	625					630					635					640
	Val	Glu	Ala	Lys	Asn	Leu	Gln	Thr	Asp	Glu	Trp	Leu	Glu	Gly	Ala	Ala
					645						650				655	
	Arg	His	Arg	Ala	Asn	Val	His	Ala	Asp	Trp	Val	His	Tyr	Trp	Gly	Gln
					660					665				670		
10	Tyr	Arg	Leu	Gly	Val	Ser	Leu	Phe	Gly	Arg	Ile	Gln	Ser	Glu	Arg	Tyr
					675					680				685		
	Tyr	Lys	Asp	Gly	Asn	Ala	Pro	Asp	Tyr	Thr	Leu	Trp	Arg	Leu	Ala	Thr
					690					695				700		
	Ser	His	Arg	Phe	Ala	His	Phe	Arg	His	Ile	Ile	Leu	Asp	Gly	Thr	Leu
15	705					710					715					720
	Gly	Ile	Asp	Asn	Leu	Phe	Asp	Tyr	Val	Asp	Asp	Arg	Pro	Met	Gly	Val
					725					730					735	
	Asn	Tyr	Ala	Thr	Val	Thr	Pro	Gly	Arg	Thr	Phe	Phe	Ala	Gln	Ile	Ala
					740					745				750		
20	Ile	Arg	Phe	Asn	Asn											
					755											

(2) INFORMATION FOR SEQ ID NO:27

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 569 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- 30 (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
- 35 (A) ORGANISM: *Porphyromonas gingivalis*
- (ix) FEATURE:
- (A) NAME/KEY: misc_feature
 - (B) LOCATION 1...569

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27

	Met	Lys	Lys	Thr	Asn	Leu	Phe	Leu	Ser	Leu	Leu	Val	Ile	Phe	Ile	Thr
	1				5					10					15	
5	Gly	Ser	Phe	Met	Thr	Ala	Cys	Ala	Gln	Lys	Ser	Lys	Thr	Asn	Lys	Leu
				20					25					30		
	Thr	Glu	Glu	Asp	Arg	Ser	Arg	Asn	Glu	Tyr	Val	Gln	Ser	Met	Asp	Val
		35						40					45			
	Leu	Ser	Asn	Ile	Ile	Gly	Asn	Val	Arg	Leu	Tyr	Phe	Val	Asp	Thr	Ile
10		50					55					60				
	Ser	Ile	Lys	His	Met	Thr	Arg	Arg	Gly	Ile	Asp	Ala	Met	Leu	Gly	Gly
	65				70					75				80		
	Leu	Asp	Pro	Tyr	Thr	Glu	Tyr	Ile	Pro	Tyr	Glu	Glu	Met	Asp	Glu	Leu
				85					90					95		
15	Lys	Leu	Met	Thr	Thr	Gly	Glu	Tyr	Ala	Gly	Val	Gly	Ala	Ile	Ile	Ser
				100					105					110		
	Gln	Arg	Pro	Asp	Ser	Ala	Val	Ile	Ile	Gln	Arg	Pro	Met	Glu	Gly	Met
		115						120					125			
	Pro	Ala	Asp	Glu	Ala	Gly	Leu	Ile	Ala	Gly	Asp	Arg	Ile	Leu	Thr	Ile
20		130					135					140				
	Asp	Gly	Lys	Asp	Phe	Arg	Lys	Ser	Thr	Thr	Pro	Lys	Val	Ser	Gln	Ala
	145				150						155				160	
	Leu	Lys	Gly	Ile	Ala	Gly	Thr	Val	Ala	Lys	Val	Thr	Val	Met	Arg	Tyr
				165					170					175		
25	Gly	Glu	Thr	Lys	Pro	Arg	Thr	Phe	Ser	Val	Lys	Arg	Gln	Lys	Val	Ile
				180					185					190		
	Met	Asn	Ser	Val	Thr	Tyr	Ser	Gly	Met	Leu	Asp	Gly	Ser	Ile	Gly	Tyr
		195						200					205			
	Ile	Arg	Leu	Asn	Asn	Phe	Thr	Asp	Lys	Ser	Ala	Glu	Glu	Val	Arg	Thr
30		210					215					220				
	Ala	Leu	Leu	Asp	Leu	Arg	Asp	Lys	Gln	Gly	Ala	Lys	Gly	Leu	Ile	Leu
	225				230					235				240		
	Asp	Leu	Arg	Gly	Asn	Gly	Gly	Gly	Leu	Met	Gln	Ala	Ala	Ile	Glu	Ile
				245					250					255		
35	Val	Asn	Leu	Phe	Val	Pro	Lys	Gly	Lys	Glu	Val	Val	Thr	Thr	Lys	Gly
				260					265					270		
	Arg	Ile	Ala	Glu	Ser	Ala	Ser	Val	Phe	Arg	Thr	Leu	Thr	Glu	Pro	Ile
		275						280					285			
	Asp	Thr	Lys	Leu	Pro	Ile	Val	Val	Leu	Ile	Asp	Gly	Gln	Ser	Ala	Ser
40		290						295					300			

Ser Ser Glu Ile Val Ala Gly Ala Leu Gln Asp Met Asp Arg Ala Val
 305 310 315 320
 Leu Met Gly Gln Lys Ser Tyr Gly Lys Gly Leu Val Gln Thr Thr Arg
 325 330 335
 5 Gln Leu Pro Tyr Asn Gly Val Ile Lys Leu Thr Thr Ala Lys Tyr Tyr
 340 345 350
 Ile Pro Ser Gly Arg Cys Ile Gln Arg Leu Asp Tyr Ser Arg Thr Asn
 355 360 365
 Arg Thr Gly Met Ala Thr Ala Ile Pro Asp Ser Leu His Lys Ile Phe
 10 370 375 380
 Tyr Thr Ala Ala Gly Arg Arg Val Glu Asp Ala Gly Gly Ile Leu Pro
 385 390 395 400
 Asp Ile Glu Val Lys Gln Asp Thr Ala Ala Thr Leu Leu Tyr Tyr Met
 405 410 415
 15 Ala Ile Asn Asn Asp Val Phe Asp Phe Val Thr Gly Tyr Val Leu Lys
 420 425 430
 His Lys Thr Ile Ala Lys Pro Glu Asp Phe Ser Ile Thr Asn Glu Asp
 435 440 445
 Tyr Ala Ala Phe Cys Lys Met Met Glu Glu Lys Lys Phe Asp Tyr Asp
 20 450 455 460
 Arg Gln Ser Gly Lys Met Leu Asp Lys Leu Glu Glu Leu Ala Lys Ile
 465 470 475 480
 Glu Gly Tyr Leu Pro Glu Ala Asn Ser Glu Leu Lys Ala Leu Arg Glu
 485 490 495
 25 Lys Leu Lys Pro Asn Leu Ser Arg Asp Leu Leu Arg Phe Lys Lys Glu
 500 505 510
 Ile Thr Asn Tyr Leu Asn Asn Glu Ile Val Thr Arg Tyr Tyr Tyr Glu
 515 520 525
 Arg Gly Ser Ile Arg Gln Ser Leu Pro Glu Asp Lys Val Val Lys Glu
 30 530 535 540
 Ala Ile Lys Leu Leu Lys Asp His Pro Glu Gln Ile Arg Gln Ile Leu
 545 550 555 560
 Ala Ala Pro Lys Ala Glu Asn Lys Gly
 565

35

(2) INFORMATION FOR SEQ ID NO:28

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 290 amino acids

(B) TYPE: amino acid

40

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5 (iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

10 (ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...290

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28

15

Met Lys Lys Leu Phe Leu Ser Leu Thr Ser Leu Val Met Val Phe Ala
1 5 10 15

Val Ala Ser Cys Asp Ile Ile Asp Lys Asp Gln Thr Leu Leu Pro Ala
20 25 30

20 Pro Thr Asn Val Thr Pro Asp Asn Pro Asp Asp Asn Pro Ser Glu Ile
35 40 45

Asp Ile Thr Gln Thr His Thr Glu Lys Tyr Val Leu Ala Glu Glu Phe
50 55 60

25 Thr Gly Gln Lys Cys Leu Asn Cys Pro Lys Gly His Arg Lys Leu Ala
65 70 75 80

Ala Leu Lys Glu Gln Tyr Gly Lys Arg Leu Thr Val Val Gly Ile His
85 90 95

Ala Gly Pro Gly Ser Leu Val Pro Pro Leu Phe Arg Thr Glu Ala Gly
100 105 110

30 Asp Ala Tyr Tyr Ser Lys Phe Ala Asn Asn Thr Pro Leu Pro Ala Leu
115 120 125

Met Val Ser Arg Lys Lys Phe Gly Ser Ser Tyr Val Tyr Asp Lys Ser
130 135 140

35 Tyr Lys Thr Trp Asp Val Pro Ile Ala Glu Gln Met Glu Gln Lys Ala
145 150 155 160

Lys Ile Asn Ile Phe Ala Val Ala Glu Tyr Thr Asp Thr Gln Lys Ile
165 170 175

Lys Val Thr Val Lys Gly Lys Ile Leu Glu Gly Asn Thr Leu Pro Lys
180 185 190

40 Ser Met Val Gln Val Tyr Leu Leu Glu Asp Lys Leu Ile Ala Pro Gln

```

195              200              205
Val Asp Gly Asn Thr Thr Val Glu Asn Tyr Glu His Asn His Val Leu
210              215              220
Arg Gly Ala Val Asn Gly Ile Trp Gly Glu Glu Phe Val Asn Leu Lys
5  225              230              235              240
Asp Tyr Leu Tyr Thr Tyr Ala Val Glu Pro Leu Ser Gly Met Ser Phe
245              250              255
Val Ala Glu Asn Tyr Ser Ile Val Ala Phe Val Tyr Asp Val Gln Thr
260              265              270
10 Phe Glu Val Tyr Asp Val Val His Val Lys Ile Asn Pro Gln Ser Asp
275              280              285
Gly Lys
290

```

15 (2) INFORMATION FOR SEQ ID NO:29

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 1017 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

25

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Porphyromonas gingivalis

(ix) FEATURE:

- 30 (A) NAME/KEY: misc_feature
 (B) LOCATION 1...1017

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29

```

35 Met Lys Arg Met Thr Leu Phe Phe Leu Cys Leu Leu Thr Ser Ile Gly
1              5              10              15
Trp Ala Met Ala Gln Asn Arg Thr Val Lys Gly Thr Val Ile Ser Ser
20              25              30
Glu Asp Asn Glu Pro Leu Ile Gly Ala Asn Val Val Val Gly Asn
40              35              40              45

```

Thr Thr Ile Gly Ala Ala Thr Asp Leu Asp Gly Asn Phe Thr Leu Ser
 50 55 60
 Val Pro Ala Asn Ala Lys Met Leu Arg Val Ser Tyr Ser Gly Met Thr
 65 70 75 80
 5 Thr Lys Glu Val Ala Ile Ala Asn Val Met Lys Ile Val Leu Asp Pro
 85 90 95
 Asp Ser Lys Val Leu Glu Gln Val Val Val Leu Gly Tyr Gly Thr Gly
 100 105 110
 Gln Lys Leu Ser Thr Val Ser Gly Ser Val Ala Lys Val Ser Ser Glu
 10 115 120 125
 Lys Leu Ala Glu Lys Pro Val Ala Asn Ile Met Asp Ala Leu Gln Gly
 130 135 140
 Gln Val Ala Gly Met Gln Val Met Thr Thr Ser Gly Asp Pro Thr Ala
 145 150 155 160
 15 Val Ala Ser Val Glu Ile His Gly Thr Gly Ser Leu Gly Ala Ser Ser
 165 170 175
 Ala Pro Leu Tyr Ile Val Asp Gly Met Gln Thr Ser Leu Asp Val Val
 180 185 190
 Ala Thr Met Asn Pro Asn Asp Phe Glu Ser Met Ser Val Leu Lys Asp
 20 195 200 205
 Ala Ser Ala Thr Ser Ile Tyr Gly Ala Arg Ala Ala Asn Gly Val Val
 210 215 220
 Phe Ile Gln Thr Lys Lys Gly Lys Met Ser Glu Arg Gly Arg Ile Thr
 225 230 235 240
 25 Phe Asn Ala Ser Tyr Gly Ile Ser Gln Ile Leu Asn Thr Lys Pro Leu
 245 250 255
 Asp Asn Met Met Thr Gly Asp Glu Leu Leu Asp Phe Gln Val Lys Ala
 260 265 270
 Gly Phe Trp Gly Asn Asn Gln Thr Val Gln Lys Val Lys Asp Met Ile
 30 275 280 285
 Leu Ala Gly Ala Glu Asp Leu Tyr Gly Asn Tyr Asp Ser Leu Lys Asp
 290 295 300
 Glu Tyr Gly Lys Thr Leu Phe Pro Val Asp Phe Asn His Asp Ala Asp
 305 310 315 320
 35 Trp Leu Lys Ala Leu Phe Lys Thr Ala Pro Thr Ser Gln Gly Asp Ile
 325 330 335
 Ser Phe Ser Gly Gly Ser Gln Gly Thr Ser Tyr Tyr Ala Ser Ile Gly
 340 345 350
 Tyr Phe Asp Gln Glu Gly Met Ala Arg Glu Pro Ala Asn Phe Lys Arg
 40 355 360 365

Tyr Ser Gly Arg Leu Asn Phe Glu Ser Arg Ile Asn Glu Trp Leu Lys
 370 375 380
 Val Gly Ala Asn Leu Ser Gly Ala Ile Ala Asn Arg Arg Ser Ala Asp
 385 390 395 400
 5 Tyr Phe Gly Lys Tyr Tyr Met Gly Ser Gly Thr Phe Gly Val Leu Thr
 405 410 415
 Met Pro Arg Tyr Tyr Asn Pro Phe Asp Val Asn Gly Asp Leu Ala Asp
 420 425 430
 Val Tyr Tyr Met Tyr Gly Ala Thr Arg Pro Ser Met Thr Glu Pro Tyr
 10 435 440 445
 Phe Ala Lys Met Arg Pro Phe Ser Ser Glu Ser His Gln Ala Asn Val
 450 455 460
 Asn Gly Phe Ala Gln Ile Thr Pro Ile Lys Gly Leu Thr Leu Lys Ala
 465 470 475 480
 15 Gln Ala Gly Val Asp Ile Thr Asn Thr Arg Thr Ser Ser Lys Arg Met
 485 490 495
 Pro Asn Asn Pro Tyr Asp Ser Thr Pro Leu Gly Glu Arg Arg Glu Arg
 500 505 510
 Ala Tyr Arg Asp Val Ser Lys Ser Phe Thr Asn Thr Ala Glu Tyr Lys
 20 515 520 525
 Phe Ser Ile Asp Glu Lys His Asp Leu Thr Ala Leu Met Gly His Glu
 530 535 540
 Tyr Ile Glu Tyr Glu Gly Asp Val Ile Gly Ala Ser Ser Lys Gly Phe
 545 550 555 560
 25 Glu Ser Asp Lys Leu Met Leu Leu Ser Gln Gly Lys Thr Gly Asn Ser
 565 570 575
 Leu Ser Leu Pro Glu His Arg Val Ala Glu Tyr Ala Tyr Leu Ser Phe
 580 585 590
 Phe Ser Arg Phe Asn Tyr Gly Phe Asp Lys Trp Met Tyr Ile Asp Phe
 30 595 600 605
 Ser Val Arg Asn Asp Gln Ser Ser Arg Phe Gly Ser Asn Asn Arg Ser
 610 615 620
 Ala Trp Phe Tyr Ser Val Gly Gly Met Phe Asp Ile Tyr Asn Lys Phe
 625 630 635 640
 35 Ile Gln Glu Ser Asn Trp Leu Ser Asp Leu Arg Leu Lys Met Ser Tyr
 645 650 655
 Gly Thr Thr Gly Asn Ser Glu Ile Gly Asn Tyr Asn His Gln Ala Leu
 660 665 670
 Val Thr Val Asn Asn Tyr Thr Glu Asp Ala Met Gly Leu Ser Ile Ser
 40 675 680 685

Thr Ala Gly Asn Pro Asp Leu Ser Trp Glu Lys Gln Ser Gln Phe Asn
 690 695 700
 Phe Gly Leu Ala Ala Gly Ala Phe Asn Asn Arg Leu Ser Ala Glu Val
 705 710 715 720
 5 Asp Phe Tyr Val Arg Thr Thr Asn Asp Met Leu Ile Asp Val Pro Met
 725 730 735
 Pro Tyr Ile Ser Gly Phe Phe Ser Gln Tyr Gln Asn Val Gly Ser Met
 740 745 750
 Lys Asn Thr Gly Val Asp Leu Ser Leu Lys Gly Thr Ile Tyr Gln Asn
 10 755 760 765
 Lys Asp Trp Asn Val Tyr Ala Ser Ala Asn Phe Asn Tyr Asn Arg Gln
 770 775 780
 Glu Ile Thr Lys Leu Phe Phe Gly Leu Asn Lys Tyr Met Leu Pro Asn
 785 790 795 800
 15 Thr Gly Thr Ile Trp Glu Ile Gly Tyr Pro Asn Ser Phe Tyr Met Ala
 805 810 815
 Glu Tyr Ala Gly Ile Asp Lys Lys Thr Gly Lys Gln Leu Trp Tyr Val
 820 825 830
 Pro Gly Gln Val Asp Ala Asp Gly Asn Lys Val Thr Thr Ser Gln Tyr
 20 835 840 845
 Ser Ala Asp Leu Glu Thr Arg Ile Asp Lys Ser Val Thr Pro Pro Ile
 850 855 860
 Thr Gly Gly Phe Ser Leu Gly Ala Ser Trp Lys Gly Leu Ser Leu Asp
 865 870 875 880
 25 Ala Asp Phe Ala Tyr Ile Val Gly Lys Trp Met Ile Asn Asn Asp Arg
 885 890 895
 Tyr Phe Thr Glu Asn Unk Gly Gly Leu Met Gln Leu Asn Lys Asp Lys
 900 905 910
 Met Leu Leu Asn Ala Trp Thr Glu Asp Asn Lys Glu Thr Asp Val Pro
 30 915 920 925
 Lys Leu Gly Gln Ser Pro Gln Phe Asp Thr His Leu Leu Glu Asn Ala
 930 935 940
 Ser Phe Leu Arg Leu Lys Asn Leu Lys Leu Thr Tyr Val Leu Pro Asn
 945 950 955 960
 35 Ser Leu Phe Ala Gly Gln Asn Val Ile Gly Gly Ala Arg Val Tyr Leu
 965 970 975
 Met Ala Arg Asn Leu Leu Thr Val Thr Lys Tyr Lys Gly Phe Asp Pro
 980 985 990
 Glu Ala Gly Gly Asn Val Gly Lys Asn Gln Tyr Pro Asn Ser Lys Gln
 40 995 1000 1005

Tyr Val Ala Gly Ile Gln Leu Ser Phe
 1010 1015

(2) INFORMATION FOR SEQ ID NO:30

5

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 811 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

10

- (ii) MOLECULE TYPE: protein

- (iii) HYPOTHETICAL: YES

15

- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Porphyromonas gingivalis

- (ix) FEATURE:

20

- (A) NAME/KEY: misc_feature
- (B) LOCATION 1...811

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30

25	Met	Arg	His	Tyr	Gly	Ile	Asn	Leu	Tyr	Glu	Leu	Lys	Arg	Leu	Ile	Glu	1	5	10	15
	Leu	Glu	Ala	Ile	Ala	Glu	Ser	Leu	Pro	Ala	Ser	Pro	Glu	Gly	Ser	Pro	20	25	30	
	Ile	Phe	Thr	Pro	Ser	Ala	Arg	Glu	Ala	Ile	Asp	Asp	Ala	Thr	Asp	Ile	35	40	45	
30	Cys	Ala	Asp	Met	Glu	Asp	Glu	Ala	Val	Ser	Pro	Val	His	Leu	Leu	Leu	50	55	60	
	Ser	Ile	Leu	Asn	Ser	Thr	Gln	Glu	Ser	Leu	Val	Gln	Lys	Ile	Phe	Met	65	70	75	80
	Lys	Gln	Gly	Ile	Lys	Tyr	Asp	Thr	Ile	Leu	Ser	Asp	Tyr	Phe	Gly	Gln	85	90	95	
35	Arg	Asn	Pro	Ser	Glu	Gly	Lys	Ser	Pro	Ser	Glu	Met	Glu	Ile	Leu	Asp	100	105	110	
	Gly	Tyr	Gln	Asp	Asn	Asp	Phe	Asp	Asp	Glu	Glu	Asp	Glu	Ser	Ser	Pro	115	120	125	
40	Pro	Ser	Gly	Asn	Ser	Gly	Thr	Gly	Gly	Gly	Ser	Gly	Asp	Ala	Pro	Glu				

	130		135		140	
	Gln Asn Thr Gly Gly Gly Asp Thr Thr Thr Thr Thr Arg Ser Gly Gly					
	145		150		155	160
	Asp Thr Pro Ala Leu Asp Thr Phe Gly Thr Asp Ile Thr Ala Met Ala					
5		165		170		175
	Ala Ala Gly Lys Leu Asp Pro Val Val Gly Arg Glu Gln Glu Ile Glu					
		180		185		190
	Arg Val Ile Gln Ile Leu Ser Arg Arg Lys Lys Asn Asn Pro Val Leu					
		195		200		205
10	Ile Gly Glu Pro Gly Val Gly Lys Ser Ala Ile Val Glu Gly Leu Ala					
		210		215		220
	Glu Arg Ile Val Asn Arg Lys Val Ser Arg Ile Leu Phe Asp Lys Arg					
		225		230		235
	Ile Ile Ser Leu Asp Leu Ala Gln Met Val Ala Gly Thr Lys Tyr Arg					
15		245		250		255
	Gly Gln Phe Glu Glu Arg Leu Lys Ala Val Leu Asp Glu Leu Lys Lys					
		260		265		270
	Asn Pro Gln Ile Ile Leu Phe Ile Asp Glu Ile His Thr Ile Val Gly					
		275		280		285
20	Ala Gly Ser Ala Ala Gly Ser Met Asp Thr Ala Asn Met Leu Lys Pro					
		290		295		300
	Ala Leu Ala Arg Gly Gln Val Gln Cys Ile Gly Ala Thr Thr Leu Asp					
		305		310		315
	Glu Tyr Arg Lys Asn Ile Glu Lys Asp Gly Ala Leu Glu Arg Arg Phe					
25		325		330		335
	Gln Lys Val Pro Ile Ala Pro Ser Thr Ala Glu Glu Thr Leu Thr Ile					
		340		345		350
	Leu Gln Asn Ile Lys Glu Lys Tyr Glu Asp Tyr His Gly Val Arg Tyr					
		355		360		365
30	Thr Asp Glu Ala Ile Lys Ala Ala Val Glu Leu Thr Asp Arg Tyr Val					
		370		375		380
	Ser Asp Arg Phe Phe Pro Asp Lys Ala Ile Asp Ala Met Asp Glu Ala					
		385		390		395
	Gly Ala Ser Val His Ile Thr Asn Val Val Ala Pro Lys Glu Ile Glu					
35		405		410		415
	Ile Leu Glu Ala Glu Leu Ala Ser Val Arg Glu Asn Lys Leu Ser Ala					
		420		425		430
	Val Lys Ala Gln Asn Tyr Glu Leu Ala Ala Ser Phe Arg Asp Gln Glu					
		435		440		445
40	Arg Arg Thr Gln Gln Gln Ile Ala Glu Glu Lys Lys Lys Trp Glu Glu					

	450		455		460	
	Gln Met Ser Lys His Arg Glu Thr Val Asp Glu Asn Val Val Ala His					
	465		470		475	480
	Val Val Ala Leu Met Thr Gly Val Pro Ala Glu Arg Leu Ser Thr Gly					
5		485		490		495
	Glu Gly Glu Arg Leu Arg Thr Met Ala Asp Asp Leu Lys Thr Lys Val					
		500		505		510
	Val Gly Gln Asp Thr Ala Ile Glu Lys Met Val His Ala Ile Gln Arg					
		515		520		525
10	Asn Arg Leu Gly Leu Arg Asn Glu Lys Lys Pro Ile Gly Ser Phe Leu					
		530		535		540
	Phe Leu Gly Pro Thr Gly Val Gly Lys Thr Tyr Leu Ala Lys Lys Leu					
		545		550		555
	Ala Glu Tyr Leu Phe Glu Asp Glu Asn Ala Met Ile Arg Val Asp Met					
15		565		570		575
	Ser Glu Tyr Met Glu Lys Phe Ser Val Ser Arg Leu Val Gly Ala Pro					
		580		585		590
	Pro Gly Tyr Val Gly Tyr Glu Glu Gly Gly Gln Leu Thr Glu Arg Val					
		595		600		605
20	Arg Arg Lys Pro Tyr Ser Val Val Leu Leu Asp Glu Ile Glu Lys Ala					
		610		615		620
	His Ala Asp Val Phe Asn Leu Leu Leu Gln Val Met Asp Glu Gly Gln					
		625		630		635
	Leu Thr Asp Ser Leu Gly Arg Arg Val Asn Phe Lys Asn Thr Val Ile					
25		645		650		655
	Ile Ile Thr Ser Asn Val Gly Thr Arg Gln Leu Lys Asp Phe Gly Gln					
		660		665		670
	Gly Ile Gly Phe Arg Ser Glu Lys Asp Glu Glu Ala Asn Lys Glu His					
		675		680		685
30	Ser Arg Ser Val Ile Gln Lys Ala Leu Asn Lys Thr Phe Ser Pro Glu					
		690		695		700
	Phe Leu Asn Arg Leu Asp Asp Ile Ile Leu Phe Asp Gln Leu Gly Lys					
		705		710		715
	Thr Glu Ile Arg Arg Met Val Asp Ile Glu Leu Lys Ala Val Leu Ala					
35		725		730		735
	Arg Ile His Arg Ala Gly Tyr Asp Leu Val Leu Thr Asp Glu Ala Lys					
		740		745		750
	Asp Val Ile Ala Thr Lys Gly Tyr Asp Leu Gln Tyr Gly Ala Arg Pro					
		755		760		765
40	Leu Lys Arg Thr Leu Gln Asn Glu Val Glu Asp Arg Leu Thr Asp Leu					

```

      770              775              780
Ile Leu Ser Gly Gln Ile Glu Lys Gly Gln Thr Leu Thr Leu Ser Ala
785              790              795              800
Arg Asp Gly Glu Ile Ile Val Gln Glu Gln Ala
5              805              810

```

(2) INFORMATION FOR SEQ ID NO:31

```

      (i) SEQUENCE CHARACTERISTICS:
10      (A) LENGTH: 293 amino acids
      (B) TYPE: amino acid
      (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: protein
15

      (iii) HYPOTHETICAL: YES

      (vi) ORIGINAL SOURCE:
      (A) ORGANISM: Porphyromonas gingivalis
20

      (ix) FEATURE:
      (A) NAME/KEY: misc_feature
      (B) LOCATION 1...293

25      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31

```

```

Met Lys Gln Asn Tyr Phe Lys Arg Val Cys Ser Leu Leu Trp Leu Val
1              5              10              15
Leu Pro Met Leu Ile Met Pro Leu Glu Val Ala Ala Gln Glu Ile Ile
30              20              25              30
Pro Asn Glu Glu Val Leu Glu Ser Leu Thr Phe Val Ala Pro Val Glu
35              40              45
Glu Thr Asp Ala Ile Glu Ala Glu Val Glu Ala Leu Gln Glu Ile Val
50              55              60
35 Ala Thr Glu Glu Ile Ala Glu Gln Ala Val Arg Ser Tyr Thr Tyr Thr
65              70              75              80
Val Tyr Arg Asp Gly Val Lys Ile Ala Ser Gly Leu Thr Glu Pro Thr
85              90              95
40 Phe Leu Asp Glu Asp Val Pro Ala Gly Glu His Thr Tyr Cys Val Glu
100              105              110

```

Val Gln Tyr Gln Gly Gly Val Ser Asp Lys Val Cys Val Asp Val Glu
 115 120 125
 Val Lys Asp Phe Lys Pro Val Thr Asn Leu Thr Gly Thr Ala Ser Asn
 130 135 140
 5 Asp Glu Val Ser Leu Asp Trp Asp Gly Val Glu Glu Lys Ala Glu Glu
 145 150 155 160
 Pro Ala Ser Asp Lys Ala Val Ser Tyr Asn Val Tyr Lys Asn Gly Thr
 165 170 175
 Leu Ile Gly Asn Thr Ala Glu Thr His Tyr Val Glu Thr Gly Val Ala
 10 180 185 190
 Asn Gly Thr Tyr Ile Tyr Glu Val Glu Val Lys Tyr Pro Asp Gly Val
 195 200 205
 Ser Pro Lys Val Ala Val Thr Val Thr Val Thr Asn Ser Ser Leu Ser
 210 215 220
 15 Asn Val Asp Gly Gln Ala Pro Tyr Thr Leu Arg Val Glu Gly Lys Lys
 225 230 235 240
 Ile Ile Ala Glu Ala His Gly Met Ile Thr Leu Tyr Asp Ile Asn Gly
 245 250 255
 Arg Thr Val Ala Val Ala Pro Asn Arg Leu Glu Tyr Met Ala Gln Thr
 20 260 265 270
 Gly Phe Tyr Ala Val Arg Phe Asp Val Gly Asn Lys His His Val Ser
 275 280 285
 Lys Ile Gln Val Arg
 290

25

(2) INFORMATION FOR SEQ ID NO:32

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 419 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

35

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

40

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...419

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32

5

```

Val Ala Phe Ile Ile Asp Leu Leu Leu Leu Cys Cys Ser Ala Phe Met
1           5           10           15
Ser Ser Cys Glu Val Ala Tyr Phe Ser Leu Lys Pro Ile Asp Leu Gln
          20           25           30
10 Asn Ile Arg Glu Arg Asn His Ser Ser Asp Ile Ala Leu Ser Asn Leu
    35           40           45
Leu Asp Asn Ser Asn Gln Leu Leu Ala Thr Ile Leu Ile Gly Asn Asn
    50           55           60
Val Ile Asn Val Ala Ile Val Ile Leu Ser Asn Tyr Ala Ile Glu Gln
15 65           70           75           80
Thr Phe Val Phe Ser Ser Pro Ile Ile Gly Phe Leu Ile Gln Thr Ile
          85           90           95
Leu Leu Thr Thr Val Leu Leu Leu Phe Gly Glu Ile Leu Pro Lys Val
          100          105          110
20 Tyr Ala Arg Lys Asn Pro Leu Gln Tyr Ser Arg Phe Ser Ala Ala Ala
    115          120          125
Met Ser Val Ile Tyr Lys Ile Leu Ser Pro Phe Ser Lys Leu Leu Val
    130          135          140
Lys Ser Thr Gly Ile Val Thr Arg Gly Ile Ser Lys Lys Lys Tyr Asp
25 145          150          155          160
Met Ser Val Asp Glu Leu Ser Lys Ala Val Ala Leu Thr Thr Thr Glu
          165          170          175
Gly Glu Pro Glu Glu Lys Glu Met Ile Asn Glu Ile Ile Lys Phe Tyr
          180          185          190
30 Asn Lys Thr Ala Cys Glu Ile Met Val Pro Arg Ile Asp Ile Val Asp
    195          200          205
Val Asp Leu Ser Trp Pro Phe Arg Lys Met Leu Asp Phe Val Val Ser
    210          215          220
Ser Gly Tyr Ser Arg Leu Pro Val Ser Glu Gly Ser Glu Asp Asn Ile
35 225          230          235          240
Lys Gly Val Ile Tyr Ile Lys Asp Leu Ile Pro His Met Asp Lys Gly
          245          250          255
Asp Glu Phe Asp Trp His Pro Leu Ile Arg Lys Ala Tyr Phe Val Pro
          260          265          270
40 Glu Asn Lys Arg Ile Asp Asp Leu Leu Glu Glu Phe Arg Ala Asn Lys

```

275 280 285
 Val His Val Ser Ile Val Val Asp Glu Phe Gly Gly Thr Cys Gly Leu
 290 295 300
 Ile Thr Met Glu Asp Ile Leu Glu Glu Ile Val Gly Glu Ile Thr Asp
 5 305 310 315 320
 Glu Tyr Asp Glu Glu Glu Leu Pro Phe Lys Val Leu Gly Asp Gly Ser
 325 330 335
 Tyr Leu Phe Glu Gly Lys Thr Ser Leu Ser Asp Val Arg His Tyr Leu
 340 345 350
 10 Asp Leu Pro Glu Asn Ala Phe Gly Glu Leu Gly Asp Glu Val Asp Thr
 355 360 365
 Leu Ser Gly Leu Phe Leu Glu Ile Lys Gln Glu Leu Pro His Val Gly
 370 375 380
 Asp Thr Ala Val Tyr Glu Pro Phe Arg Phe Gln Val Thr Gln Met Asp
 15 385 390 395 400
 Lys Arg Arg Ile Ile Glu Ile Lys Ile Phe Pro Phe Glu Arg Thr Trp
 405 410 415
 Glu Val Glu

20

(2) INFORMATION FOR SEQ ID NO:33

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 853 amino acids

25

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

30

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

35

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION 1...853

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33

40

Met Leu Pro Leu Pro Tyr Arg Tyr Ala Lys Thr Glu His Leu Phe Leu
 1 5 10 15
 Ala Lys Gly Tyr Cys Lys Asn Pro Ile Thr Asn Ile Ile Ile Leu Phe
 20 25 30
 5 Met Lys Lys Lys Asn Phe Leu Leu Leu Gly Ile Phe Val Ala Leu Leu
 35 40 45
 Thr Phe Ile Gly Ser Met Gln Ala Gln Gln Ala Lys Asp Tyr Phe Asn
 50 55 60
 Phe Asp Glu Arg Gly Glu Ala Tyr Phe Ser Phe Lys Val Pro Asp Arg
 10 65 70 75 80
 Ala Val Leu Gln Glu Leu Ala Leu Ile Met Ser Ile Asp Glu Phe Asp
 85 90 95
 Pro Val Thr Asn Glu Ala Ile Ala Tyr Ala Ser Glu Glu Glu Phe Glu
 100 105 110
 15 Ala Phe Leu Arg Tyr Gly Leu Lys Pro Thr Phe Leu Thr Pro Pro Ser
 115 120 125
 Met Gln Arg Ala Val Glu Met Phe Asp Tyr Arg Ser Gly Glu Lys Tyr
 130 135 140
 Glu Trp Asn Ala Tyr Pro Thr Tyr Glu Ala Tyr Ile Ser Met Met Glu
 20 145 150 155 160
 Glu Phe Gln Thr Lys Tyr Pro Ser Leu Cys Thr Thr Ser Val Ile Gly
 165 170 175
 Lys Ser Val Lys Asp Arg Lys Leu Met Ile Cys Lys Leu Thr Ser Ser
 180 185 190
 25 Ala Asn Thr Gly Lys Lys Pro Arg Val Leu Tyr Thr Ser Thr Met His
 195 200 205
 Gly Asp Glu Thr Thr Gly Tyr Val Val Leu Leu Arg Leu Ile Asp His
 210 215 220
 Leu Leu Ser Asn Tyr Glu Ser Asp Pro Arg Ile Lys Asn Ile Leu Asp
 30 225 230 235 240
 Lys Thr Glu Val Trp Ile Cys Pro Leu Thr Asn Pro Asp Gly Ala Tyr
 245 250 255
 Arg Ala Gly Asn His Thr Val Gln Gly Ala Thr Arg Tyr Asn Ala Asn
 260 265 270
 35 Asn Val Asp Leu Asn Arg Asn Phe Lys Asp Asp Val Ala Gly Asp His
 275 280 285
 Pro Asp Gly Lys Pro Trp Gln Pro Glu Ala Thr Ala Phe Met Asp Leu
 290 295 300
 Glu Gly Asn Thr Ser Phe Val Leu Gly Ala Asn Ile His Gly Gly Thr
 40 305 310 315 320

Glu Val Val Asn Tyr Pro Trp Asp Asn Lys Lys Glu Arg His Ala Asp
 325 330 335
 Asp Glu Trp Tyr Lys Leu Ile Ser Arg Asn Tyr Ala Ala Ala Cys Gln
 340 345 350
 5 Ser Ile Ser Ala Ser Tyr Met Thr Ser Glu Thr Asn Ser Gly Ile Ile
 355 360 365
 Asn Gly Ser Asp Trp Tyr Val Ile Arg Gly Ser Arg Gln Asp Asn Ala
 370 375 380
 Asn Tyr Phe His Arg Leu Arg Glu Ile Thr Leu Glu Ile Ser Asn Thr
 10 385 390 395 400
 Lys Leu Val Pro Ala Ser Gln Leu Pro Lys Tyr Trp Asn Leu Asn Lys
 405 410 415
 Glu Ser Leu Leu Ala Leu Ile Glu Glu Ser Leu Tyr Gly Ile His Gly
 420 425 430
 15 Thr Val Thr Ser Ala Ala Asn Gly Gln Pro Leu Lys Cys Gln Ile Leu
 435 440 445
 Ile Glu Asn His Asp Lys Arg Asn Ser Asp Val Tyr Ser Asp Ala Thr
 450 455 460
 Thr Gly Tyr Tyr Val Arg Pro Ile Lys Ala Gly Thr Tyr Thr Val Lys
 20 465 470 475 480
 Tyr Lys Ala Glu Gly Tyr Pro Glu Ala Thr Arg Thr Ile Thr Ile Lys
 485 490 495
 Asp Lys Glu Thr Val Ile Met Asp Ile Ala Leu Gly Asn Ser Val Pro
 500 505 510
 25 Leu Pro Val Pro Asp Phe Thr Ala Ser Pro Met Thr Ile Ser Val Gly
 515 520 525
 Glu Ser Val Gln Phe Gln Asp Gln Thr Thr Asn Asn Pro Thr Asn Trp
 530 535 540
 Glu Trp Thr Phe Glu Gly Gly Gln Pro Ala Met Ser Thr Glu Gln Asn
 30 545 550 555 560
 Pro Leu Val Ser Tyr Ser His Pro Gly Gln Tyr Asp Val Thr Leu Lys
 565 570 575
 Val Trp Asn Ala Ser Gly Ser Asn Thr Ile Thr Lys Glu Lys Phe Ile
 580 585 590
 35 Thr Val Asn Ala Val Met Pro Val Ala Glu Phe Val Gly Thr Pro Thr
 595 600 605
 Glu Ile Glu Glu Gly Gln Thr Val Ser Phe Gln Asn Gln Ser Thr Asn
 610 615 620
 Ala Thr Asn Tyr Val Trp Ile Phe Asp Gly Gly Thr Pro Ala Thr Ser
 40 625 630 635 640

Glu Asp Glu Asn Pro Thr Val Leu Tyr Ser Lys Ala Gly Gln Tyr Asp
 645 650 655
 Val Thr Leu Lys Ala Ile Ser Ala Ser Gly Glu Thr Val Lys Thr Lys
 660 665 670
 5 Glu Lys Tyr Ile Thr Val Lys Lys Ala Pro Val Pro Ala Pro Val Ala
 675 680 685
 Asp Phe Glu Gly Thr Pro Arg Lys Val Lys Lys Gly Glu Thr Val Thr
 690 695 700
 Phe Lys Asp Leu Ser Thr Asn Asn Pro Thr Ser Trp Leu Trp Val Phe
 10 705 710 715 720
 Glu Gly Gly Ser Pro Ala Thr Ser Thr Glu Gln Asn Pro Val Val Thr
 725 730 735
 Tyr Asn Glu Thr Gly Lys Tyr Asp Val Gln Leu Thr Ala Thr Asn Glu
 740 745 750
 15 Gly Gly Ser Asn Val Lys Lys Ala Glu Asp Tyr Ile Glu Val Ile Leu
 755 760 765
 Asp Asp Ser Val Glu Asp Ile Val Ala Gln Thr Gly Ile Val Ile Arg
 770 775 780
 Pro Gln Asn Gly Thr Lys Gln Ile Leu Ile Glu Ala Asn Ala Ala Ile
 20 785 790 795 800
 Lys Ala Ile Val Leu Tyr Asp Ile Asn Gly Arg Val Val Leu Lys Thr
 805 810 815
 Thr Pro Asn Gln Leu Arg Ser Thr Val Asp Leu Ser Ile Leu Pro Glu
 820 825 830
 25 Gly Ile Tyr Thr Ile Asn Ile Lys Thr Glu Lys Ser Ala Arg Thr Glu
 835 840 845
 Lys Ile His Ile Gly
 850